NMR CHARACTERIZATION OF LOCAL UNFOLDING IN *E. COLI* ADENYLATE KINASE.

by

Jeremy A. Anderson

A dissertation submitted to Johns Hopkins University in conformity with the requirements for the degree of Doctor of Philosophy.

Baltimore, Maryland

January, 2017

Abstract

Life depends on proper function of biological macromolecules such as proteins. However, the application of knowledge about proteins and other macromolecules suffers from a lack of understanding of basic properties. This is further evident in the study of enzymes that catalyze chemical reactions. Conformational fluctuations between the ground state and one or more intermediate states have emerged as a strong candidate for the missing link in the canonical structure-function paradigm. Such states have shown to be important in enzyme function, allosteric signaling, and ligand binding. These processes are difficult to study due to low populations of one or more conformations and their transient nature. In this study, we investigated a locally unfolded (LU), intermediate state in *E. coli* adenylate kinase (AK). The LU state was stabilized using entropy-enhancing mutations that increase the degeneracy of unfolded conformations while maintaining the original ground-state structure. The population of the LU state was tuned to varying degrees, giving access to different equilibrium populations. Nuclear magnetic resonance was used to probe the structural characteristics of the LU state and measure conformational dynamics between the ground and LU states. Chemical shift differences observed directly and inferred from dynamics matched one another and are consistent with an LU intermediate conformation for all variants studied. This study directly characterizes a functionally important locally-unfolded, intermediate conformation in AK and exemplifies the ability of entropy-enhancing mutations to study unfolded conformations in proteins.

Advisor: Vincent J. Hilser, PhD

Reader: Doug Barrick, PhD

Committee: Juliette Lecomte, PhD

Bertrand Garcia-Moreno E., PhD

Kalina Hristova, PhD

Preface

The completion of my Ph.D. could not have been possible without the support, care, and love of an army of people. There are many points in my life where I have been guided to the next step. The stern recommendations of Mrs. Martine and Mr. Bednar and the inspiration of Mr. Scheuer at Owen-With High School were instrumental in getting me to attend college. At the University of Minnesota, Duluth Dr. Jonathan Maps taught me how to think about physics and Dr. Anne Hinderliter showed me how it fit into biology. Drs. Grant Anderson and Thomas Bastian gave me the opportunity to do research for the first time and welcomed my input as an equal.

In graduate school, I was met by the loving personality of Ranice Crosby. She, along with the rest of the biophysics staff not only made life easier but kept me grounded. My fellow students in biology and biophysics have provided support both scientifically and personally that made my time at Hopkins bearable and more enjoyable. As time goes on I appreciate more and more the guidance of the biophysics faculty and the work they put into maintaining a great program and training graduate students. In particular, I am grateful for the patience of my thesis committee: Juliette Lecomte, Doug Barrick, and Bertrand Garcia-Moreno E. Lastly, I must that Dr. Ananya Majumdar who has not only taught me everything I know about NMR but has been a great friend and mentor.

The driving force in my development as a scientist has been Dr. Vincent J. Hilser and the other members of the Hilser Lab. Past and present members of the lab have fostered an excellent community for learning and doing science including: Dr. James Wrabl, Dr. W. Austin Elam, Dr. Jing Li, Andrew Martens, Jordan White, James Rives, thanks Obama, Joe Rehfus, Emily Grasso, Min Hyung Cho, and especially Dr. Hesam Motlagh and Alex Chin.

My journey if life would not have been possible without the love and support of my family. First and foremost, the guidance of my wife on a daily basis helps me to stay on the beam and continually strive to do better. She has put up with me as a person and the requirements of graduate school. Her skill in special education has been paramount to both. My brother Shawn who has supported and nudged me at

every step along the way. My dad, who may have given me too much freedom at times but let me chase whatever dreams struck me. And finally, my grandparents, Kristian and Margaret Rasmussen. They are the kindest, most caring individuals I have ever met. They have taught me the most valuable lessons in life and their love has carried me when I couldn't carry myself. I can find no words to describe my gratitude. To them I dedicate this thesis.

Table of Contents

List of Tables

List of Figures

Chapter 1 – Introduction

1.1 Proteins, Enzymes, & Temperature

A common set of macromolecules perform the exquisitely complex set of interactions that orchestrate life. Proteins drive a majority of these interactions: for processing genetic information, for utilizing energy and materials from food, for providing cell structure and machinery to make new cells, and for transferring information within and between cells.[1] Therefore, it is no surprise that much effort has gone into studying protein chemistry and biophysics, although, as with all molecular biology, this a relatively recent endeavor. Basic discoveries like the determination that a protein is a linear chain of amino acids and that enzymes catalyzing chemical reactions are proteins were not verified until the first quarter of the 20th century.[2] However, huge advancements in biophysics have been made since. Some of these include the elucidation of complex folds on the nanometer scale or the application of thermodynamics to describe stability and quantitative measurements of protein folding.[3-6] New methodologies like computational simulations, bioinformatics, and experimental techniques like single-molecule experiments continue to improve our understanding of proteins.[7]

Much of the work in protein biophysics has focused on globular, folded proteins. This has provided organizing principles about many concepts like protein structure, folding, binding and catalysis.[8-11] In general globular proteins fold to form stable, three-dimensional structures encoded by the sequence of amino acids.[6] Folding is usually two-state and cooperative. Many proteins have moderate stability due to large, competing enthalpies and entropies of folding. An unstructured, denatured ensemble of conformations folds rapidly, with few to no intermediates, to a unique folded conformation. This observation clashed with a theoretical consideration of protein folding, termed Levinthal's Paradox: if an unfolded protein samples all possible conformations as it folds it would take extremely long to reach the native folded state. This paradox was later reconciled with the formation of energy landscape theory and

protein folding funnels as models for protein folding.[9, 12-14] Proteins populate a Boltzmann distribution of conformations based on their free energy. The unfolded state is unstable and has many different conformations. As local, native structure forms, the conformations accessible to the polypeptide are biased towards the native, folded conformation. The protein is gradually stabilized during this process preventing return to the unfolded state further biasing conformational space. This sequential stabilization process has also been observed experimentally in hydrogen-deuterium mass spectrometry experiments that measured sequence-specific folding and unfolding of cytochrome C.[15-16] These models provide a means to move beyond a static, structural view of proteins and investigate conformational dynamics.

Since the advent of structural biology over fifty years ago, it has been proposed that many proteins require movement or conformational flexibility to function. Many experiments have verified the importance of dynamics in dictating different aspects of protein function.[17-21] For instance, the ability of hemoglobin to cooperatively bind and release oxygen over small changes in concentration depends on conformational changes within the protein.[22] This phenomenon, termed allostery or action at a distance, occurs in many proteins. Structural interpretations of these phenomena paint a picture of a finely engineered system, consisting of specific pathways of interactions. However, as structural information accumulated and protein design became possible it was realized that structure does not fully determine function and that phenomena like allostery can be the result of conformational fluctuations.[19, 23-25] Ensemble-based models of proteins that take into account local stability and fluctuations from the ground state structure Intermediates in protein folding have also been observed and seem to involve local fluctuations involving unfolding. Intrinsically disordered proteins, or IDPs, also highlight the importance of conformational ensembles and fluctuations between states. This class of proteins remain unfolded, accessing many different conformations, and sometimes fold upon binding with interaction partners.[19, 26] Investigations of fluctuations between conformations has been difficult but is rising as methodologies become generally accessible. Nonetheless, the prevalence of high-energy conformational states is not known.

Enzymes illustrate the role of conformational fluctuations in protein function. In terms of transition-state theory the activation energy of a reaction determines the rate. Fluctuations amongst an ensemble of protein conformations are thought play a role in determining the activation energy. Examples of protein motions that affect enzyme function exist but it remains controversial whether conformational fluctuations are necessary for function.[27-28] Conformational dynamics in enzymes have been theorized to be crucial to the chemical reactions that these proteins catalyze; however, some argue that mere stabilization of the transition state is sufficient for enzyme function.[29-32] A less contentious point is that conformational fluctuations can determine the rate-limiting step in catalysis. Gaps in the basic understanding of protein dynamics may be preventing the design of rationally engineered enzymes, which have long been a goal yet remain unrealized.[33] The limited success of both computational and high throughput experimental approaches in designing modified or de novo enzymes highlights the complexity underlying enzyme function.[34-35] Although only a scaffold stabilizing the transition state may be required for catalysis, protein dynamics clearly modulate enzyme activity. [36]

Enzyme temperature adaptation also demonstrates that conformational flexibility contributes to function. Cold-adapted, or psychrophilic, enzymes generally have faster reaction rates and lower binding affinities compared to warm-adapted homologues when measured at the same temperature.[37-43] These effects are due to favorable decreases in enthalpies and unfavorable decreases in entropies of the reaction resulting in similar free energies of activation but shallower temperature dependencies.[38] The changes in energetic parameters are consistent with an expansion of the transition-state ensemble of conformations to include low affinity states that require structural rearrangements upon binding. Additionally, activities of warm-adapted enzymes are maximum at or near the melting temperature, losing activity as proteins unfold while cold enzymes lose activity at temperatures well below the unfolding transition of the protein.[38, 44] Loss of activity below the global unfolding transition is consistent with lower thermal stability of regions where conformational fluctuations are important for catalysis. Lastly, active sites of homologues are highly conserved whereas surface residues and loops commonly differ with many of these mutations appearing to increase the conformational entropy.[38-40, 43, 45] These examples are consistent with an increase in conformational entropy and an expansion of the ensemble of available conformations in cold-adapted

proteins. Such mutations compensate for reduced thermal energy, resulting in increased activity over warm-adapted homologues at low temperatures.[45] This suggests that changes in the ensemble of available conformations are capable of fine-tuning enzyme activity while maintaining the ground state structure and exist as an additional feature upon the scaffold of the ground state structure of globular proteins. Furthermore, mutations that increase conformational entropy in cold-adapted proteins should maintain the native state structure of the protein. The studies here are an attempt to gain a quantitative understanding of the nature of conformational flexibility in proteins and its role in enzyme function.

1.2 *E. coli* Adenylate Kinase

This thesis focuses on conformational fluctuations in the enzyme adenylate kinase (AK) from *Escherichia coli*. AK is a 23.5 KDa (214 residue) protein made up of three domains as depicted in Figure 1.1: a central CORE domain (residues 1-27, 73-119, 161-214) and two insertion domains: the AMP binding domain (AMPbd residues 28-72) and the LID domain (residues 120-160). This enzyme is required as it catalyzes the reaction $ATP + AMP \Leftrightarrow 2 \cdot ADP$, which maintains energy homeostasis in the cell. Crystal structures have been solved in the presence and absence of the bisubstrate inhibitor AP5A (P1,P5-di(adenosine-5′) pentaphosphate), showing the LID and AMPbd closing over the active site upon binding. From structural analysis, this conformational change is due to rigid body rotation of both insertion domains about "hinges" between the "open," apo form to a "closed," bound form.[46-47]

Many experimental and computational studies have focused on the opening/closing of the LID and AMPbd. The fact that both of these domains close over the active site both in the presence and absence of ligand is well established.[48-53] In addition to capturing the end states of the opening/closing transition, crystal structures have been solved that place the LID domain in between the two states.[54-57] Shortly after this time-resolved FRET studies found evidence for step-wise, single domain closure in the presence and absence of ligand. Single-molecule FRET later found that AK is in equilibrium between fully open and fully closed states even in the apo form and that ligand shifted the equilibrium to the closed form by

increasing the closing rate of the LID.[58] Single-molecule FRET also found that the closed state of the enzyme was accessed in the presence and absence of ligand. However, computational studies have found that closing of both the LID and AMPbd are unstable in the absence of ligand.[59-60] Dual-bead single molecule tweezers experiments with tethers attached to the LID and AMPbd measured distance changes consistent with cooperative closing of the insertion domains upon binding of AP5A.[52] In competition assays with ATP and ADP the same experiment can resolve single domain closure and extract binding affinities for these ligands. Molecular dynamics studies investigating the effect of different ligands on the equilibrium between the open and closed states also found independent closure of single domains upon binding.[61] Other computational studies have found that opening and closing of the LID and AMPbd is cooperative.[62-63] However, the opening and closing rates have been found to depend significantly on the force field used in molecular dynamics simulations.[64] Eximer formation has been observed for pyrene molecules attached to each the LID and AMPbd in the presence of ATP, ADP, AMP and AP5A and indicate closure of both domains upon binding.[65] This study also found that closure was strongest when ligand binds simultaneously to both domains as is the case with either AP5A or ADP. Time-resolved FRET and NMR experiments have observed increases in the population of the closed LID in the presence of stabilizing osmolyte concomitant with increased binding affinity for ligand.[66-67] In a compelling study, paramagnetic relaxation enhancement NMR data places multiple residues of the LID domain within 30 angstroms of the AMPbd in the absence of ligands, which strongly supports closure of both domains.[51] The native equilibrium between open and closed is not clear since attachment of probes and DNA handles can influence the equilibrium in unknown ways. From the above studies it seems clear that AK is in equilibrium between the open and closed conformations in the presence and absence of ligand. Furthermore, the closed conformation appears to be less stable and has a higher affinity for ligands.

Opening of the LID domain is believed to be the rate-limiting step in catalysis. This mechanism was first proposed after the elucidation of both the open and closed structures of the protein and described by rigid body motions involving up to 8 hinge points within the protein.[47] Many NMR studies have investigated conformational dynamics in the presence of different ligands, including AP5A, as well as during enzymatic turnover with a mixture of ATP, AMP, or ADP. Conformational dynamics have been

observed for residues in these hinge regions on the order of the catalytic rate of the enzyme.[48, 50] Alternatively, these data could be the result of dynamic binding; ring current shifts, phosphate groups, and extensive hydrogen bond networks could all influence the data.[68] The hinge regions have also been found to be evolutionarily conserved; however, the most conserved residues either interact with AP5A directly or stabilize structural elements containing residues that do. A single mutation that increases catalytic activity has been found in a hinge region near the LID domain; however, it also breaks a salt bridge between the LID domain (residue E170) and the AMPbd (residue L58).[67] This salt bridge, one of many connecting the LID and AMPbd domains, has been computationally observed to determine the enthalpic barrier to domain opening.[55] Experimentally swapping the hinge regions between mesophilic and thermophilic AKs results in no change in enzymatic activity. [49] However, swapping the entire LID and AMPbd results in thermophilic activity in the mesophilic protein and vice versa, which questions the role of LID domain in exclusively determining the catalytic rate. An investigation into the role of in catalysis found that Mg^{2+} facilitated both phosphoryl transfer and opening of the LID domain due to interaction spanning the active site and the LID.[54]

An equally valid interpretation of these data are that the dynamics of the AMPbd are rate limiting. This is consistent with catalytic data collected in this laboratory (Harry Saavedra, unpublished work). Computational work suggested local unfolding as a means to alleviate strain in the closed state of the LID and AMPbd, in turn assisting the transition to the open state.[69] Of note is the fact that the LID domain is evolutionarily expendable, that is many organisms contain indels within the LID region or have nearly no LID domain at all. The variability in the LID domain over evolutionary timescales is not understood, nor are the functional consequences of truncating the domain. Interestingly, the LID domain of some thermophilic AKs have been shown to coordinate Zn^{2+} ions similar to zinc beta-ribbon domains.[70] These studies show an incomplete picture of the AK mechanism. Hinge regions may be important in determining catalytic activity; however, other regions of the protein, particularly the LID and AMPbd, appear responsible for fine-tuning the enzyme. How this occurs is not evident from the experiments discussed above. Therefore, this laboratory has set out to identify high-energy conformations that may contribute to AK function.

1.3 Entropy-enhancing Mutations & AK

High energy conformations that differ from crystal structures of AK have not been identified. It may be that there are no other accessible states or that populations are too low to observe. Theoretical descriptions of the energetic landscape of proteins predicts the existence of high-energy, disordered states that have also been observed in hydrogen-deuterium exchange experiments.[16, 71-74] Enzyme evolution suggests that increasing conformational entropy can promote these states. Therefore, an entropy-enhancing mutational approach was developed to modulate the equilibrium populations of high-energy, unfolded conformations that may be important for tuning enzyme function. Valine to glycine mutations were made at solvent-exposed residues at a distance greater than 8 angstroms from the active site of the enzyme. Glycine residues lack a β carbon and therefore have greater phi/psi accessibility. A valine to glycine mutation increases the number of unfolded conformations by about four-fold leading to a stabilization of roughly 1 kcal/mol at 25°C.[75-79] Such mutations mimic the increased flexibility observed in cold adapted enzymes while leaving the original ground-state structure largely unaffected. This is in contrast to common mutational strategies that affect internal packing or electrostatic effects. The two mutations studied here are both in the LID domain at positions V135 and V142 as shown in Figure 1.3. It appears that this strategy is limited to probing unfolded states, however, such mutations could stabilize transition state ensembles between structured conformations that involve partial unfolding. This view is in agreement with local unfolding as a mechanism to reduce strain or frustration in proteins, which has been proposed to drive opening of the LID and AMPbd.[69, 80]

Previous work in this laboratory has shown evidence for local unfolding and changes in enzyme function upon entropy-enhancing mutations.[81-83] LID domain variants showed reduced binding affinity for the bisubstrate analog AP5A at higher temperatures. Nonlinear decreases in binding enthalpy, as seen in Figure 1.2 A, were observed at increasing temperature and required a 3-state binding model. This behavior corresponded to an increase in the LU population at higher temperatures as shown by the population diagram in Figure 1.2 C. The enthalpy and heat capacity change of this conformational transition is

consistent with local unfolding and match those from theoretical calculations of LID unfolding. Loss of LID peak intensity for V142G at high temperatures in HSQC spectra corroborated this result. No differences between the crystal structures of the V135G variant and WT AK were observed, consistent with a preservation of ground state structure. Furthermore, single LID mutations have no effect on the catalytic activity of the enzyme over a wide temperature range as shown in Figure 1.2 B. This is at odds with the determination that LID domain dynamics determine the catalytic rate. It could be that unfolding does not affect LID opening; however this seems unlikely and is inconsistent with computational studies.[69] These previous results provide strong evidence for local unfolding in the LID domain; however, direct observation of the LU conformation and the kinetics of exchange have not been determined.

1.4 Outline of this Thesis

This thesis describes the direct characterization of a locally unfolded, intermediate state in adenylate kinase using nuclear magnetic resonance spectroscopy. In the WT protein, this intermediate state is experimentally unobservable and has evaded study for the last 20 years. However, the entropy-enhancing mutation strategy shifts equilibrium populations from fully folded to locally unfolded, bringing intermediate states into focus. In HSQC experiments the V142G variant remains folded whereas the V135G/V142G LID is entirely unfolded. The V135G variant shows a mixture of both folded and LU populations in slow exchange. Structural characterization of the LU conformation from these variants is described in Chapter 2. Kinetics of this transition were quantified using NMR relaxation dispersion experiments. The V142G and V135G/V142G variants both have skewed populations with only one conformation clearly visible. However, as explained in Chapter 3, both variants show conformational dynamics to another state. A combination of relaxation dispersion and ZZ exchange NMR experiments were used to characterize the slow exchange process in the V135G variant as detailed in Chapter 4. Furthermore, this chapter explores the nature of the LU transition and attempts to reconcile all of the data in the context of two-state and multi-state exchange. While the quantification of intermediate-state

conformations of proteins is relatively new, many studies have used the NMR techniques employed here to gain information on such states.[36, 84-91] However, rational modulation of populations through an enhancement of dynamics has not been achieved previously. Therefore, in addition to extending the knowledge of the conformational landscape of AK and proteins in general this study exemplifies the beneficial pairing of entropy-enhancing mutations and NMR experiments to probe dynamics. This mutational strategy should prove to be generally applicable to any protein system with dynamically driven, energetically accessible intermediate states, which could prove to be a useful tool at a time when such systems are moving to the spotlight.

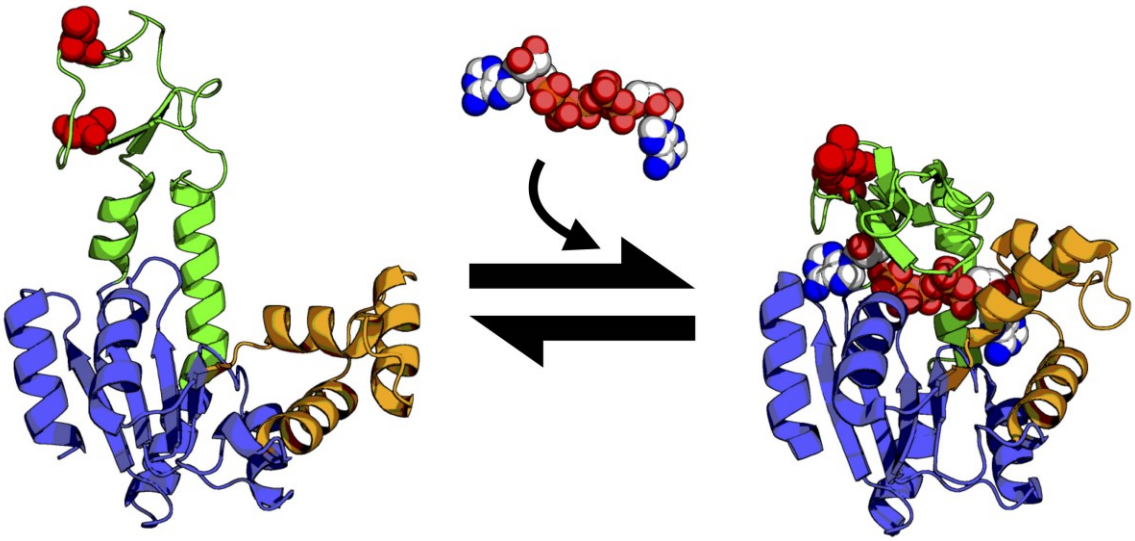

Figure 1.1 Crystal structures of AK

Open and closed crystal structures of AK (PDB 4AKE & 1AKE) with the LID domain in green, AMPbd in orange, and CORE in blue, showing large conformational change upon binding of the bisubstrate inhibitor AP5A.[47, 92] The positions of entropy-enhancing mutations, V142 and V135, are shown in red spheres

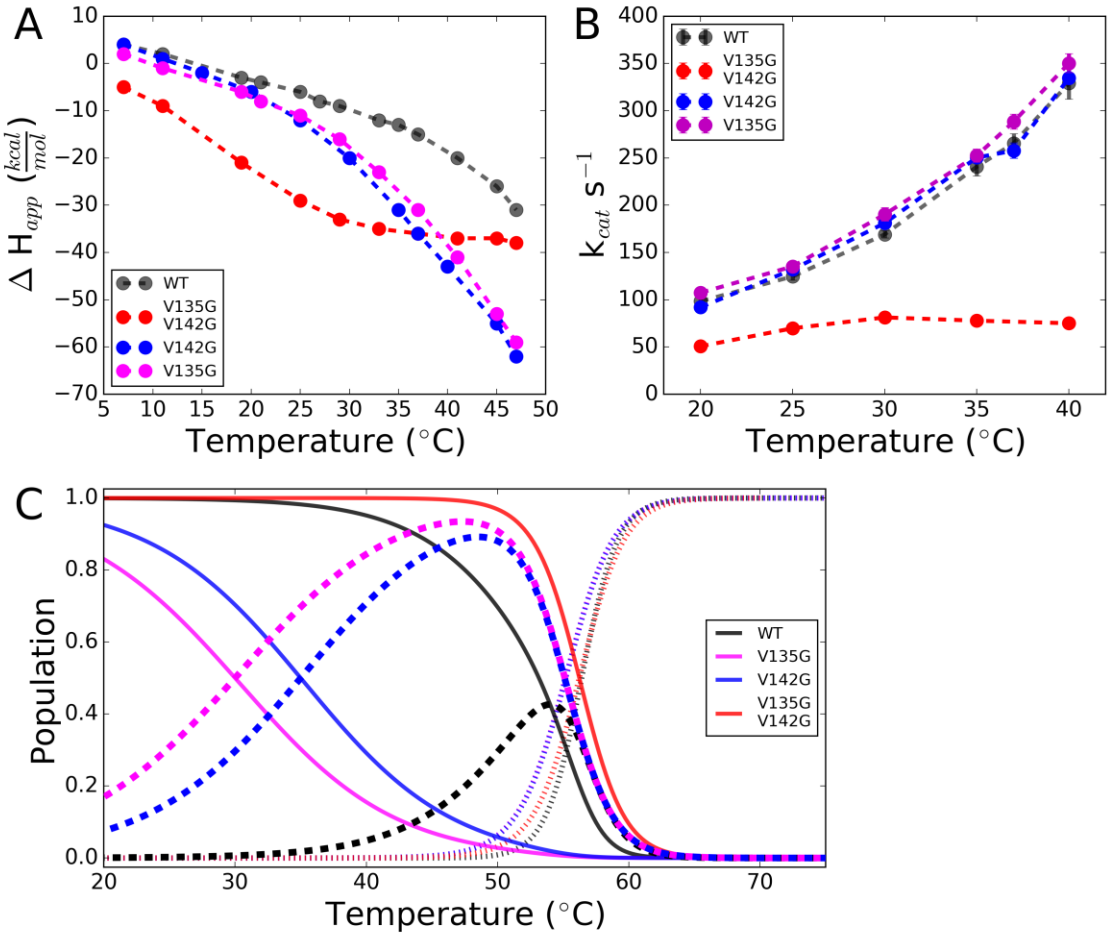

Figure 1.2 Binding, Catalysis, and Thermodynamics of AK Variants

A. Apparent enthalpy of binding from ITC experiments of variants with the bisubstrate analog AP5A. The nonlinear increase in negative enthalpy at high temperatures is due to the contribution of increasing population of locally unfolded molecules. The double variant had no discernable folded population even at low temperatures. B. Catalytic activity for the reverse phosphotransferase reaction $2 \cdot AD \rightarrow ATP + AMP$ shows no change in single variants compared to the WT activity while the double variant has severely reduced activity. C. Populations of the folded (solid line), unfolded (dotted line), and LU (dashed line) states determined from differential scanning calorimetry for WT, V142G, and V135G show that entropy-enhancing mutations stabilize the LU state. The solid red line represents the LU population in the V135G/V142G variant as the folded conformation is not observable. Data courtesy of Harry Saavedra.

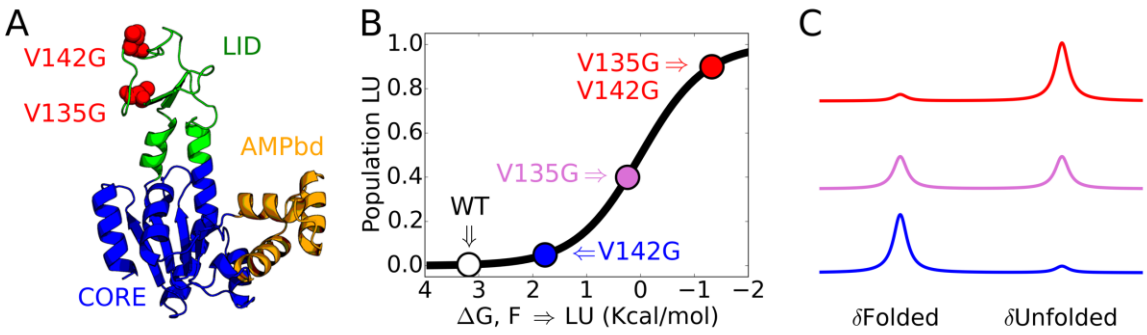

Figure 1.3 Effect of Entropy-Enhancing Mutations

A. Mutations in the LID domain at positions V135 and/or V142 give access to points along a local unfolding curve at 19°C (B) as the LU state is stabilized resulting in a shift in populations observed in NMR experiments from folded to unfolded resonances as depicted in C. Populations of the LU state in B and C are estimated from observed populations in NMR experiments, which do not agree with calorimetric values for the V135G variant.

Chapter 2 – Structure of the LU conformation in AK variants

2.1 Abstract

The LID domain of adenylate kinase (AK) populates an unfolded conformation that is enhanced with entropy-enhancing mutations. Previous thermodynamic and NMR experiments on single LID variants presented evidence for local unfolding, however, the LU state was not verified structurally.[81] In this study, NMR data unequivocally show that the LID domain locally unfolds upon mutation. Furthermore, the population of the LU state can be tuned by utilizing multiple mutation sites in the LID. This approach allows access to different equilibrium populations: the folded baseline with V142G, the LU baseline with V135G/V142G, and equal mixture of folded and LU populations with V135G. Observed chemical shift changes are consistent with unfolding and are conserved amongst all variants demonstrating that variants promote the same unfolded conformation. Furthermore, these results highlight the generality of entropy-enhancing mutations and confirms that they have little to no effect on the ground state structure.

2.2 Introduction

High-energy intermediate-state conformations of proteins are difficult to characterize structurally since they are lowly populated and transient. Entropy-enhancing mutations can stabilize these states so they can be studied experimentally. Previous experiments established that entropy-enhancing mutations in the LID domain of AK promote a binding incompetent state consistent with a locally unfolded conformation. However, structural information about this state was lacking. Here, the structural characteristics of the binding incompetent state are explored using NMR. NMR allows for site-specific interrogation of structural differences between variants. While relatively simple, HSQC experiments provide a wealth of information on the local chemical environment of a given backbone amide and provide

a structural "fingerprint" of the protein.[93] Differences in chemical shifts between variants inform on the type of conformational change that occurs and the structure of the intermediate-state conformation.[94] Furthermore, backbone amides undergoing conformational exchange on slow timescales manifest as two separate peaks in the HSQC experiment allowing for a direct measurement of the chemical difference between the conformations. Chemical shift differences measured between slow exchange pairs can then be compared to differences observed between other variants. Using this approach, we show that: no structural changes are present between the WT protein and the V142G variant, two mutations shift V135G/V142G LID peaks to an unfolded conformation, and V135G is in slow exchange between both of these conformations. These studies show that the binding-incompetent, intermediate state previously identified is a conformation with an unfolded LID domain. Furthermore, multiple entropy-enhancing mutations bring this state into focus highlighting the generality of this mutation scheme.

2.3 Materials & Methods

Preparation of Entropy-Enhancing Variants

Double Colony Selection

AK constructs were expressed using an ampicillin resistant pJ414 vector in BL21 DE3 pLysS *E. coli* cells. Transformants went through a double colony selection procedure to ensure cell growth and robust expression in deuterated minimal media.[95] Transformed cells were plated on a 70% D_2O, LB/AMP plate and 9 colonies used to inoculate 9-1mL LB/AMP cultures in 100% D_2O. These cultures were then shaken at 250 rpm and 37°C for 7-10 hours. Concurrently, a 70% D_2O master plate was divided into 9 sections, and spotted with each corresponding colony ~10 times by touching a pipette tip to the colony on the transformation plate and repeatedly touching the corresponding section of the master plate without touching the original colony in between. The master plate was incubated at 37°C for 24 hours and stored at 4°C until double colony selection was complete. Upon reaching an optical density at 600 nm (OD_{600}) of

1.0 (mid log phase) the LB cultures were used to inoculate 9-1 mL, 100% D_2O MOPS minimal medium cultures with 1% unlabeled glucose to an initial OD_{600} of 0.05-0.1 and shaken at 250 rpm and 37°C. After reaching an optical density of 0.6-1.0, MOPS cultures were induced with 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for ~10 hours at 37°C. After expression, 250 μL of cells were collected and boiled in Laemmli sample buffer in preparation for SDS-PAGE. The amount of cell lysate loaded in the gel was normalized by the final OD_{600} of each colony. The colony with the greatest expression from the gel was used to inoculate a 70% D_2O LB culture using a single colony from the master plate. This culture was grown to an OD of ~1 and 100 μL of a $1:10^5$ dilution spread on a 70 % D_2O LB/AMP plate. This plate was used to repeat the above colony selection process once more. Finally, the colony with the greatest expression after the second round of selection was used to inoculate a 70% D_2O LB culture, grown to an OD_{600} of ~1, and used to make glycerol stocks stored at 80°C.

Expression & Purification

After colony selection ampicillin plates were streaked from glycerol stocks and a single colony was used to inoculate a 500 mL LB culture. Initially LB cultures were shaken at 150 rpm and 25°C for 14 hours (overnight) then at 250 rpm and 37°C until an OD_{600} of 1.0 was reached. The LB culture was then spun down at 3580 g for 15 minutes and resuspended in 250 mL of 100% D_2O MOPS minimal media at 0.4% $^2H/^{12}C$ or $^2H/^{13}C$ glucose depending on the isotope required. The MOPS culture was grown for one hour at 250 rpm to assimilate cells to the new conditions and then induced with 1 mM IPTG for 8 hours. Cultures were then spun down at 3580 g for 15 minutes at 4°C. The cell pellet was collected and frozen at -80 °C. To begin purification the cell pellet was thawed and sonicated in three volumes of 50 mM TRIS, 0.1 M NaCl, 2 mM EDTA, 2 mM DTT, pH 8.0, and a protease inhibitor cocktail tablet (Roche cOmplete ULTRA Tablets, Mini, EASYpack, reference 05892970001) for 15 minutes. Cell lysate was then spun down at 15000 rpm and 4°C for one hour. The supernatant was filtered three times at 0.22 μm, diluted with two volumes of filtered Cibracon Blue A buffer (50 mM TRIS, 0.1 mM EDTA, 1 mM dithiothreitol, and pH 7.5) and loaded onto a 50 mL Cibracon Blue affinity (for ATP binders) column pre-equilibrated with A buffer. The column was then washed with two column volumes of A Buffer and eluted with a linear

gradient from 0-100% B buffer (A buffer plus 1 M NaCl) over 200 mL while collecting 5 mL fractions. Peak fractions from the chromatogram were run on an SDS-PAGE gel to check that AK was present (molecular weight of 23.5 KDa) and inspect purity. Note that owing to the generality of the binding to the Cibracon Blue column samples were purified in batches from 1 L of expression medium or less. This prevented overloading of the column with unwanted proteins, which can lead to contamination of the end product. Fractions containing AK were then pooled and dialyzed against 4 L of the Resource Q A buffer (20 mM TRIS, 0.1 mM EDTA, 1 mM DTT, pH 8.0) in 3 KDa molecular weight dialysis tubing for 16 hours at 4°C. Dialyzed protein was loaded onto a 75 mL quarternary amine, Resource Q anion exchange column pre-equilibrated with A buffer. The column was washed with two column volumes of A Buffer and eluted with a linear gradient from 0-100% Resource Q B buffer (A buffer with 0.1 M NaCl) over 300 mL collecting 5 mL fractions. Peak fractions were run on SDS-PAGE gels to assess purity. Pure fractions containing AK were pooled and dialyzed against 2 L of the AK NMR buffer (50 mM HEPES, 20 mM $MgCl_2$, 5 mM TCEP, pH 7.0) for 16 hours at 4°C in a 3 kDa molecular weight dialysis cassette. The protein was then concentrated by centrifugation at 25°C and 3580 g to 1 mM, spiked with 10% D_2O, degassed for 30 minutes, and loaded into a Shigemi NMR tube. Using these protocols highly deuterated samples were routinely acquired at concentrations of at least 12 mg/L of expression medium.

NMR Experiments, Data Processing, and Analysis

Theory

The ZZ exchange measures exchange phenomenon occurring on the slow NMR timescale where the exchange rate constant is much lower than the chemical shift difference between the two conformational states. This process manifests as two peaks in the HSQC spectrum, one for each conformation, as can be seen in Figure 2.11. The ZZ exchange experiment can assign residues in slow exchange and quantify the kinetics of the exchange process, which is expanded upon in Chapter 4 below. In the HSQC version of the experiment a mixing time is added that allows molecules to exchange conformations after labeling magnetization with the ^{15}N chemical shift and before detecting the ^{1}H

chemical shift. This results in peaks with chemical shifts from both conformations. For instance, if a molecule begins in state A and changes conformation to state B during the mixing time a peak will appear at the nitrogen shift of A and the proton shift of B, $^{1}H_{B}$-$^{15}N_{A}$. Cross peaks are dependent on the mixing time and can be used to quantify the kinetics of the dynamic process as discussed below. A 2D $^{1}H/^{1}H$ experiment can be used for quick verification of slow exchange as is shown in Figure 2.8. Furthermore, 3D experiments can be used to further separate peaks and assign slow exchange pairs.

Experiment & Analysis

^{1}H-^{15}N TROSY HSQC spectra were collected with 16 transients, a 70-90 ms acquisition time in the nitrogen dimension, and a 2.5 s delay between transients at both 14.7 and 18.8 T field strengths at 19°C for all variants.[96] HNCA, HNCACB, HNCO, HNCACO, and ^{15}N-edited NOESY experiments were used to assign backbone resonances of the unfolded conformation of the LID domain in the V135G/V142G variant at 19°C. All spectra were processed using NMRPipe{Delaglio, 1995 #460} and in house scripts; data were visualized, peaks assigned, and peak heights/volumes calculated and exported using the CCPNMR analysis program.[97]

Referencing differences between exported peak lists, due to slight temperature differences, were corrected based on median chemical shift differences as compared to WT spectra observed in both proton and nitrogen. This was done for all residues excluding 100-180. Median differences were calculated for all spectra as compared to the WT spectrum using in-house python scripts. Combined chemical shift differences were calculated based on the following equation[98]:

$$\Delta\delta = \left\{(\delta^{1}H)^{2} + \left[\frac{(\delta^{15}N)}{6.51}\right]^{2}\right\}^{1/2} \quad (1)$$

2.4 Results & Discussion

Structural Similarity between WT & V142G

The LID domain of V142G remains folded at low temperatures but unfolds at higher temperatures. At 19 °C the V142G variant retains a folded structure similar to the WT protein as can be seen in an overlay of the two HSQC spectra shown in Figure 2.1. Many peaks in the WT spectrum do not change in V142G; however, small peaks appear within the amide proton region between 7.5-8.5 ppm. These additional peaks have been present in all sample preparations, at least three, and stable for over one year suggesting they are not contamination degradation. Additional information about minor peaks could not be deduced owing to weak intensity and was not included in subsequent analyses. Backbone assignments were transferred from WT to V142G and verified using through-space NOE measurements. Differences in chemical shifts were then quantified and are shown in Figure 2.2. The differences observed are quite small across the entire protein and the largest perturbations occur in the LID domain near the mutation site as might be expected. These data demonstrate that the structure of the LID domain in V142G at 19°C is largely folded and similar to the WT protein, matching expectations based on thermodynamic data. At 25°C most LID peaks in V142G disappear but remain in the WT spectrum as seen in Figure 2.3. At higher temperatures, minor peaks in the V142G spectrum do not intensify and new peaks are not observed while some folded peaks remain. These results suggest that most LID residues in V142G move into intermediate exchange and convalesce at higher temperatures rather than just shifting peak intensity to the LU population.

LID Unfolding V135G/V142G

Mutating valines 135 and 142 to glycine in AK completely unfolds the LID domain. This can be seen in an overlay of the V142G HSQC spectrum over V135G/V142G at 19°C shown in Figure 2.4. A significant number of peaks appear or intensify in the center of the amide proton spectrum in the

V135G/V142G variant. The reverse overlay (V135G/V142 over V142G) in Figure 2.5 shows LID peaks that are present in V142G that disappear in the double variant. Assignment of backbone resonances in V135G/V142G verified that LID domain residues had indeed collapsed into the center of the amide proton spectrum as depicted by the pseudo HSQC in Figure 2.6. This region is a hallmark of unfolded proteins and demonstrates that V135G/V142G LID residues are largely unfolded.[99] Combined chemical shift differences, shown in Figure 2.7, demonstrate that the structural change compared to the WT is large and confined to the LID domain. These structural observations agree with thermodynamic experiments that were unable to detect any folded LID population even at low temperatures (unpublished data). The N-terminus near glycine 10 also shows a structural rearrangement. These residues pack against a helix at the base of the LID domain and are presumably sensitive to unfolding of the LID. These residues make up the p-loop that is essential for ATP binding and are highly conserved. Disruption of p-loop structure could explain the reduced binding affinity of LID variants. The data from V135G/V142G demonstrate that two entropy-enhancing mutations are sufficient to destabilize the LID domain even at low temperatures making the LU conformation predominant as depicted in Figure 1.3.

V135G in Slow Exchange between F and LU

Mutation of valine 135 to glycine moves the LID domain into slow exchange between folded and unfolded conformations. Many peaks were present in the HSQC spectrum of V135G that overlaid with LID assignments for both V142G and V135G/V142G while no peaks were present at some folded LID positions. This overlay suggested that most V135G LID residues were in slow exchange between the folded, ground and LU conformations. To verify exchange between these peaks ZZ exchange NMR experiments were performed. In these experiments, a mixing time allows amino acids to switch populations, from fold to LU or vice versa, during the course of the experiment. Mixing populations correlates the two exchanging peaks resulting in cross peak intensity at the intersection of the two peaks in a 2D experiment like an HSQC, i.e. peaks at positions $^{1}H_{A}{}^{15}N_{B}$ and $^{1}H_{B}{}^{15}N_{A}$. The extent of slow exchange in the V135G variant can be easily seen in the 2D $^{1}H/^{1}H$ ZZ exchange spectrum shown in Figure 2.8. In

this spectrum, diagonal peaks represent the 1D amide proton spectrum and off-diagonal elements correlate peaks belonging to the same proton from the same amino acid. Two points are immediate from the figure, first there are a large number of residues in slow exchange, and second the proton chemical shift difference is large for many residues signifying a dramatic structural rearrangement. Slow exchange pairs were assigned using additional 2D and 3D ZZ exchange experiments. One peak from each pair was denoted either folded or LU for peaks that overlaid best with V142G or V135G/V142G, respectively. The pseudo-HSQC in Figure 2.9 shows the slow exchange pairs as blue and red circles for the folded and LU populations, respectively, with gray wedges connecting peaks recorded from the same ^{15}N and ^{1}H nuclei belonging to the same amino acid in a similar fashion as Figure 2.6. All LU peaks collapse into the center of the amide proton spectrum, characteristic of an unfolded protein. Furthermore, the combined chemical shift difference between the two populations in Figure 2.10 shows large changes localized to the LID domain. The fact that V135G stabilizes the LU state to a greater degree than V142G is surprising and in disagreement with thermodynamic data that predicted an LU population of 15% at 19°C. A large part of this discrepancy is most likely due to differences in transverse relaxation rates. The intrinsic ^{15}N relaxation rates for LU peaks are around 15 s^{-1} while the ground state peaks have rates around 30 s^{-1}, which can lead to a nearly two-fold increase in peak height for the same area when considering only the nitrogen dimension. Furthermore, additional peak broadening due to conformational exchange on intermediate timescales could reduce the height of one or both peaks.

All variant comparison

LID mutations have subtle and interesting effects in other regions of the protein. Chemical shift differences were calculated between WT and V142G, V135G/V142G, and the folded set of V135G peaks to assess changes in the structure of the folded conformation. Separate proton and nitrogen chemical shift differences are shown in Figure 2.11. Interestingly, there are significant changes outside of the LID domain in addition to the aforementioned changes around residue 10 in the p-loop. For instance, the largest proton chemical shift difference in V135G occurs at phenylalanine 81 located in the third beta strand in

CORE domain. The same residues show differences in both the V142G and V135G/V142G variants in the same direction as V135G. However, the trend in the magnitude of the differences in both proton and nitrogen is such that V135G has the greatest effect followed by V135G/V142G and finally V142G. This same trend is observed in other regions of the protein including near residues 2, 103, and to a lesser extent 26, and 190. All of these residues are located at a common edge of the beta sheet that makes up the central CORE domain, as shown in Figure 2.12, suggesting structural rearrangement of this region upon unfolding of the LID domain. However, the trend of chemical shift changes in this region is in contrast to the observed trend in LU populations for LID residues where V135G/V142G results in the greatest change in populations. This indicates that while related to local unfolding, chemical shift changes within the beta sheet of the CORE are not directly correlated to the population of the LU state.

A comparison of LID peaks suggests that all variants undergo the same local unfolding transition. The magnitude of the combined chemical shifts presented in Figure 2.7 and Figure 2.10 are quite similar between the V135G and V135G/V142G variants as are the direction of the changes as seen in the pseudo-HSQC spectra in Figure 2.6 and Figure 2.9. The chemical shift differences between V142G and V135G/V142G were compared directly to those between slow exchange pairs in V135G. Figure 2.13 shows excellent agreement between the two sets of chemical shift differences in both proton and nitrogen even with all common LID residues included. This correlation demonstrates that entropy-enhancing mutations in the LID domain have the same effect in all three variants and all variants access the same transition.

2.5 Conclusions

The structural data presented here show unequivocally that the LID domain locally unfolds and entropy-enhancing mutations stabilize the LU conformation. The V142G variant shows no structural evidence of local unfolding but the disappearance of peaks at higher temperatures suggests conformational exchange in the LID. Two mutations in the LID shift the observed equilibrium entirely to the LU

conformation. Remarkably, LID residues in V135G are in slow exchange between the folded and LU conformations. Although the results for V142G and V135G/V142G are consistent with previous thermodynamic data, V135G appears to stabilize the LU conformation to a greater extent than expected. Greater flexibility of locally unfolded LID residues, resulting in lower relaxation rates and greater peak intensities, could explain the difference in the observed equilibrium population by NMR versus calorimetric results. Additionally, conserved, tandem proline residues at positions 139 and 140 could restrict flexibility in the unfolded state and potentially reduce the effect of the V142G mutation. Multiple exchange processes could weaken peak intensity of one population and result in observed populations in NMR experiments that are different from thermodynamic results. Multi-state exchange will be discussed in greater detail in Chapter 4 below. Variants with different populations of the LU state provide an opportunity to characterize the LU transition from multiple perspectives strengthening the analysis. It is remarkable that such small mutations can produce such a large effect in the LID domain of AK. This can only be possible if the stability of the LID domain is poised sufficiently close to the local unfolding transition as depicted in Figure 1.3. Local instability has been suggested as a method for reducing protein strain and for tuning enzymes.[38, 43, 100-101] Entropy-enhancing mutations could be used to test these hypotheses and provide a useful tool for investigating the effects of protein dynamics in other systems.

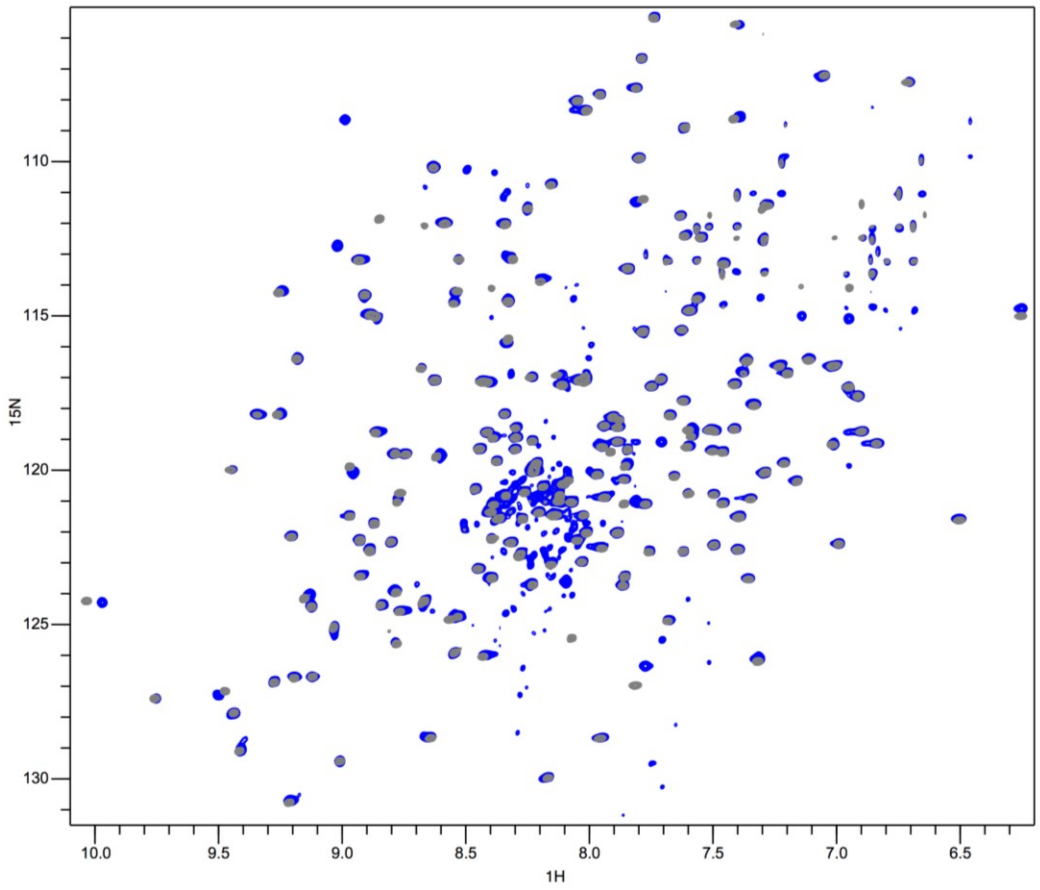

Figure 2.1 WT over V142G HSQC Overlay

TROSY HSQC spectra of WT AK (grey) over V142G (blue) at 19 °C and a pH of 7.0 shows that most peaks in both spectra overlap with one another. Reproducible minor peaks in the middle of the amide proton spectrum may be evidence of the LU conformation or sample degradation.

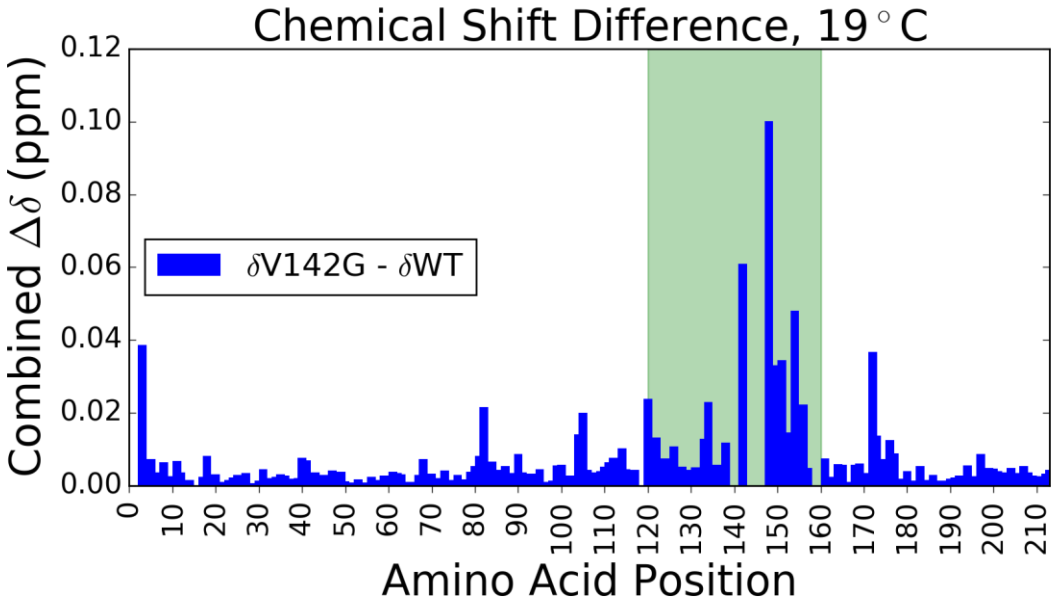

Figure 2.2 V142G Combined Chemical Shift Difference

The combined ^{1}H ^{15}N chemical shift difference between WT and V142G is largest in the LID domain, shaded in green. However, differences are small throughout the protein demonstrating that the structure of V142G is similar to that of the WT protein.

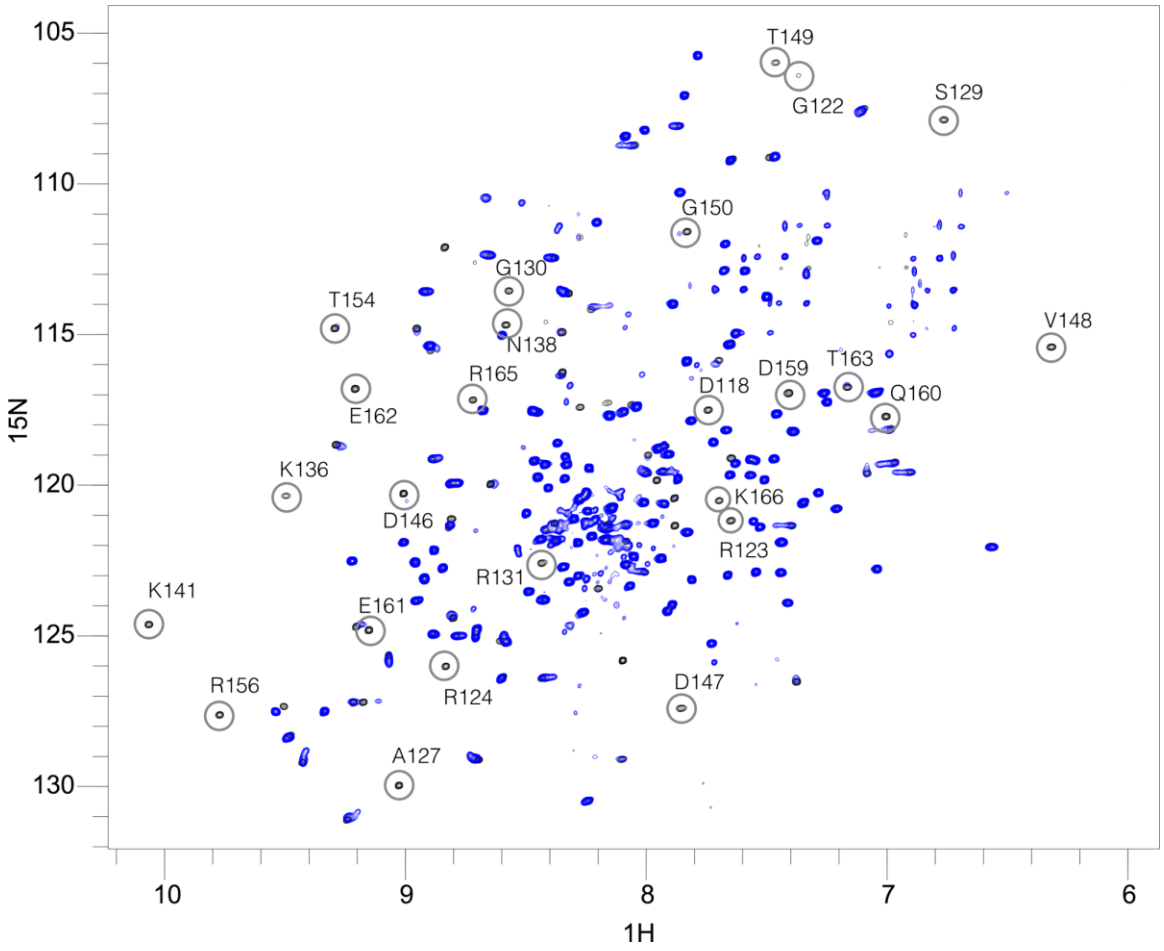

Figure 2.3 V142G over WT HSQC Overlay 25℃

The TROSY HSQC spectrum of V142G (blue) overlaid on top of the WT (black) spectrum at 25 °C and a pH of 7.0 shows that many LID peaks disappear in V142G (circled, labeled peaks) while they remain in the WT. These data suggest that LID residues are moving from slow to intermediate exchange in V142G as some folded LID peaks remain and the estimated LU population is only 15%.

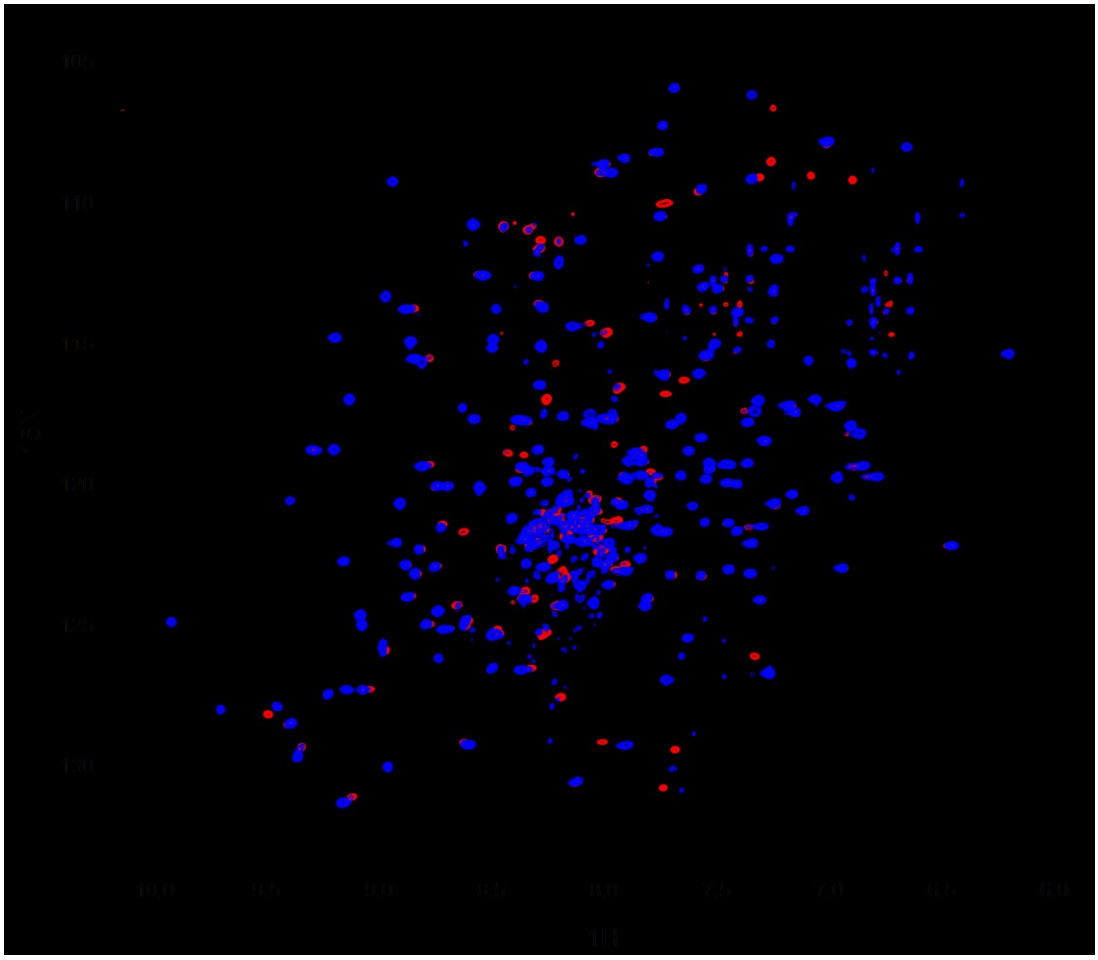

Figure 2.4 V142G over V135G/V142G HSQC Overlay 19℃

The TROSY HSQC spectrum of V135G/V142G (red) under that of V142G (blue) both collected at 19 °C and a pH of 7.0 shows increased peak intensity and new peaks in the middle of the amide proton spectrum, consistent with an increased LU population.

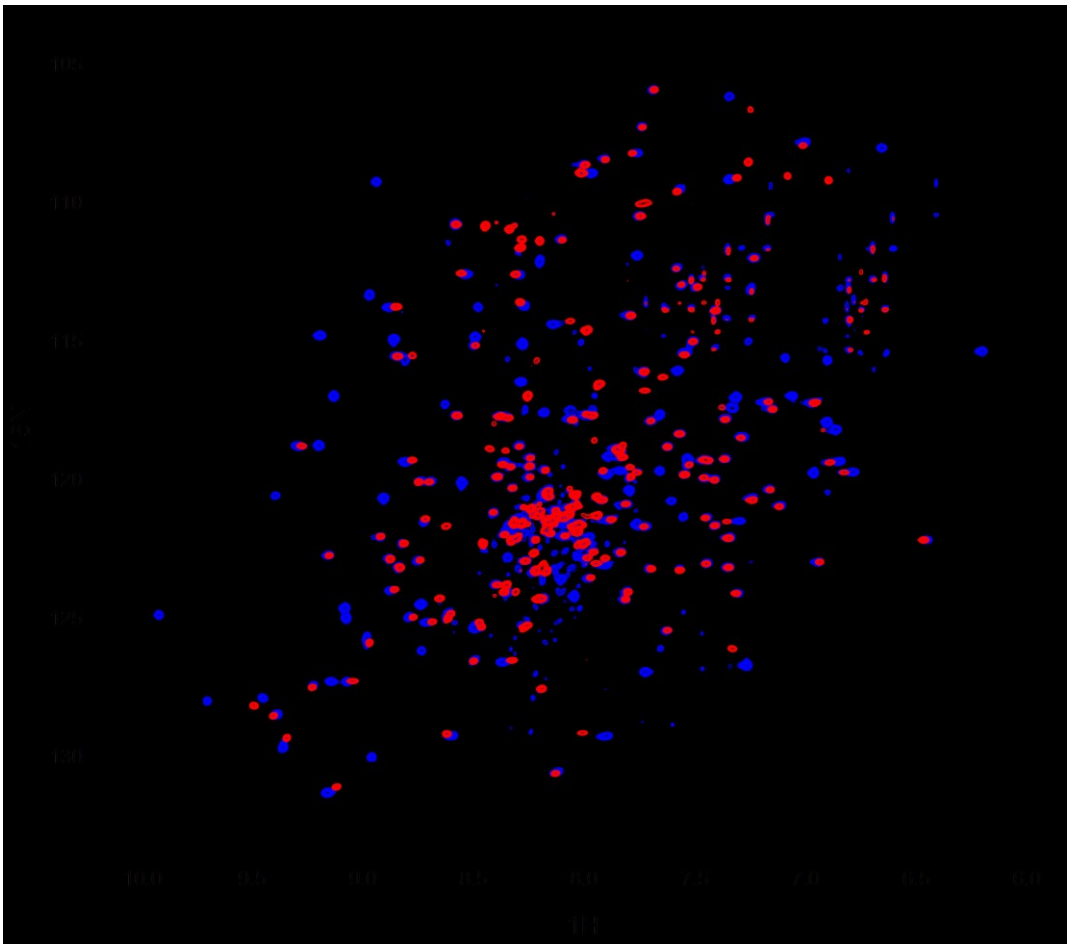

Figure 2.5 V135G/V142G over V142G HSQC Overlay 19℃

The TROSY HSQC overlay of V135G/V142G (red) over V142G (blue) both collected at 19 °C and a pH of 7.0 shows a loss of peak intensity upon a second mutation for many residues similar to Figure 2.3.

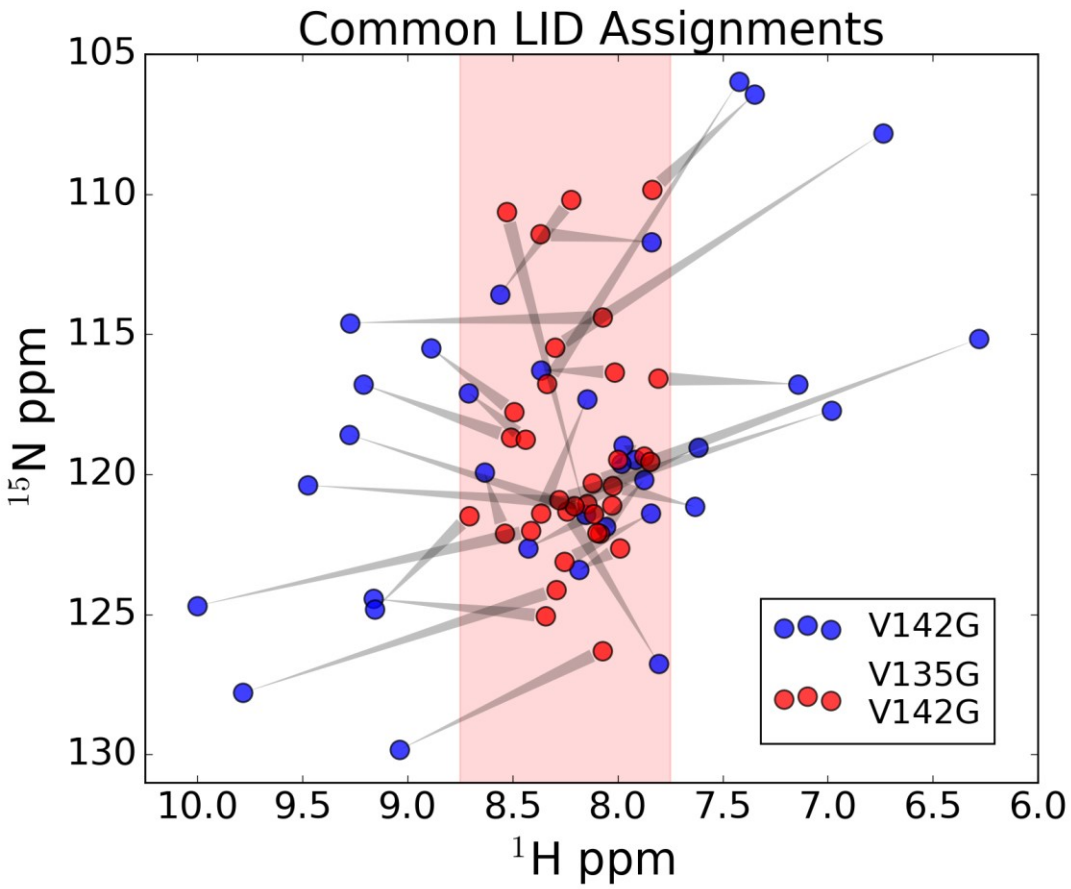

Figure 2.6 V142G V135G/V142G Pseudo-HSQC

A pseudo-HSQC of common peak assignments between V142G (blue) and V135G/V142G (red) at 19 °C with grey wedges connecting peaks for the same amino acid. This plot shows that upon a second mutation in the LID domain peaks collapse into the center of the amide proton spectrum, shaded in red, consistent with an unfolded state.

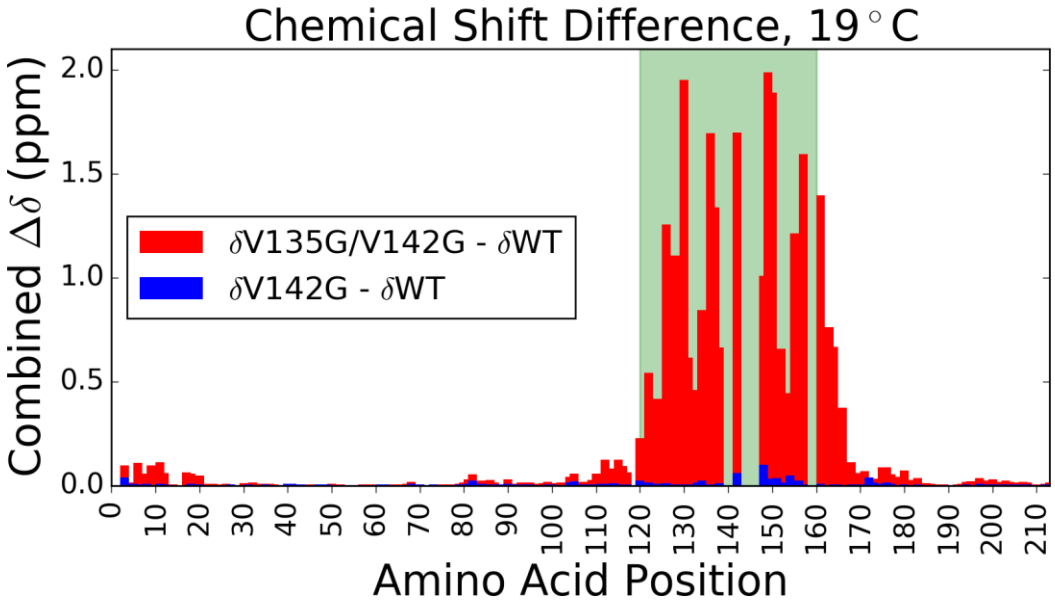

Figure 2.7 V135G/V142G Combined Chemical Shift Difference

Combined ^{1}H-^{15}N chemical shift differences in the V135G/V142G variant, shown in red, are confined to the LID domain (green) demonstrating that unfolding is a local effect. V142G chemical shift difference plotted as in Figure 2.2.

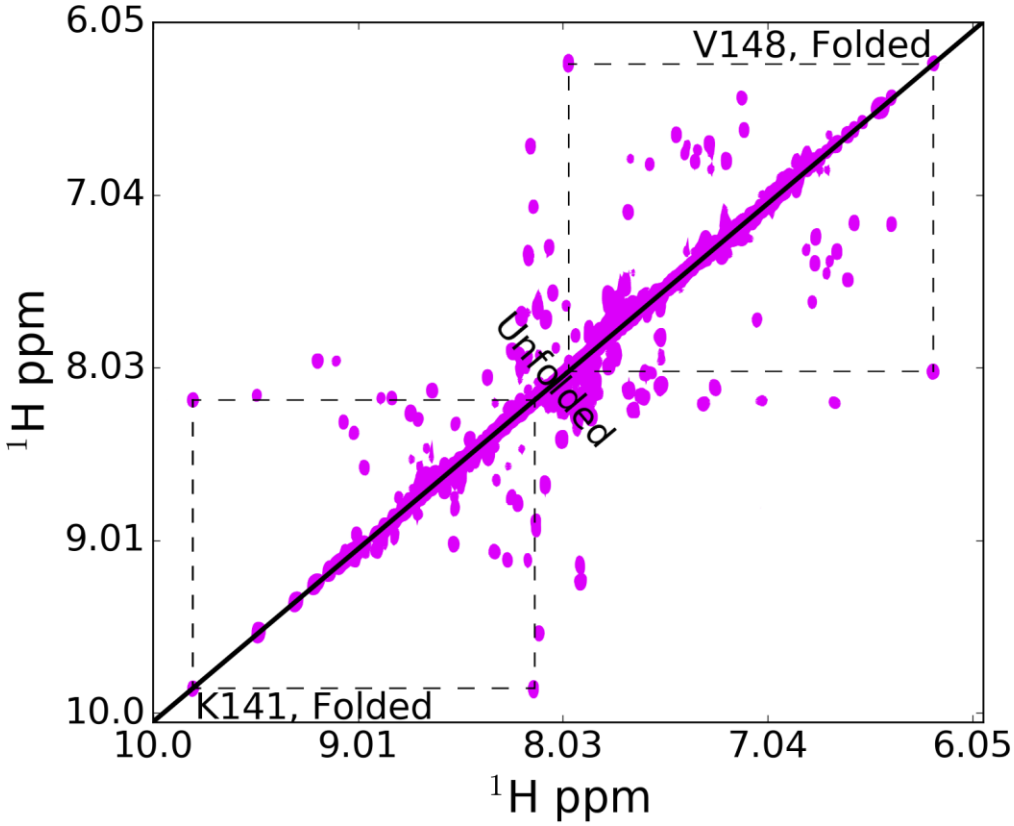

Figure 2.8 V135G ZZ ^{1}H/^{1}H Spectrum

The ^{1}H/^{1}H ZZ exchange spectrum collected on V135G at 22°C shows direct evidence of slow exchange between amide proton (diagonal) peaks that belong to the same amino acid in folded and unfolded conformations. Diagonal peaks for residues in exchange can be connected by cross peaks as shown for K141 and V148.

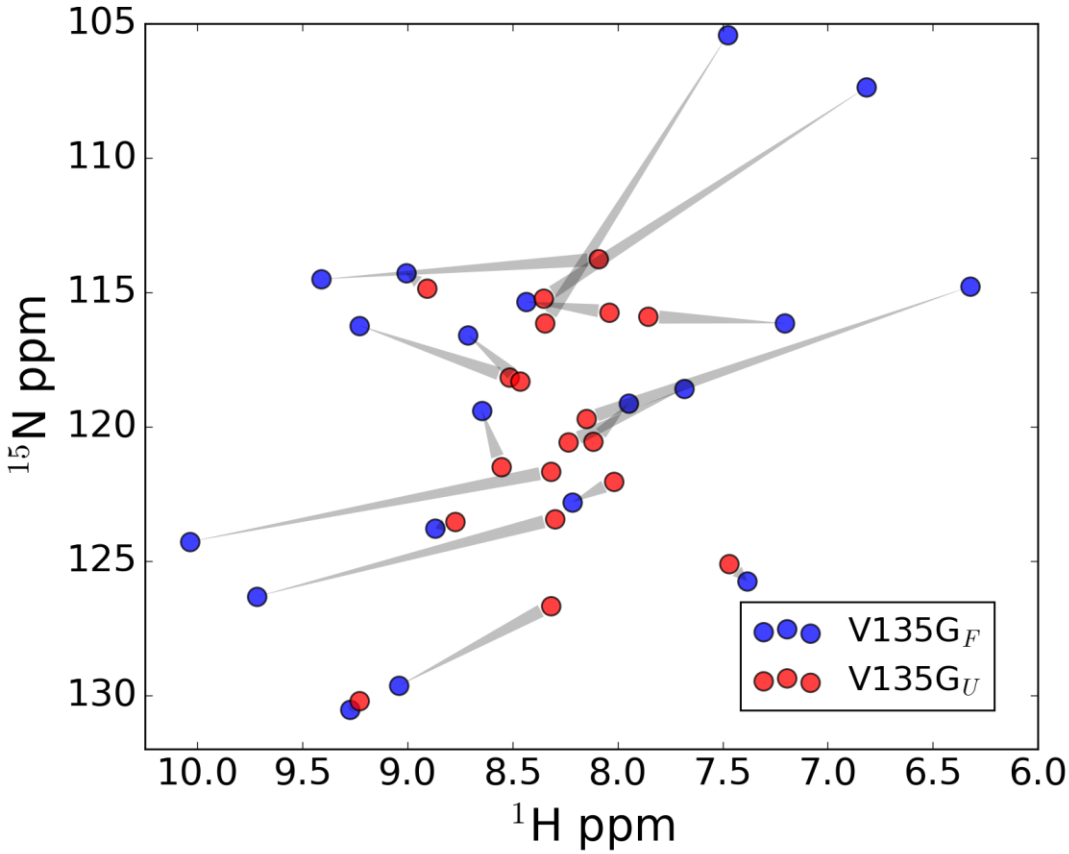

Figure 2.9 V135G Pseudo-HSQC

A pseudo HSQC of assigned slow exchange pairs in V135G at 19 °C shows large structural changes similar to Figure 2.6 with peaks in blue representing folded positions that overlay with WT and V142G and peaks in red representing unfolded positions that overlay with V135G/V42G.

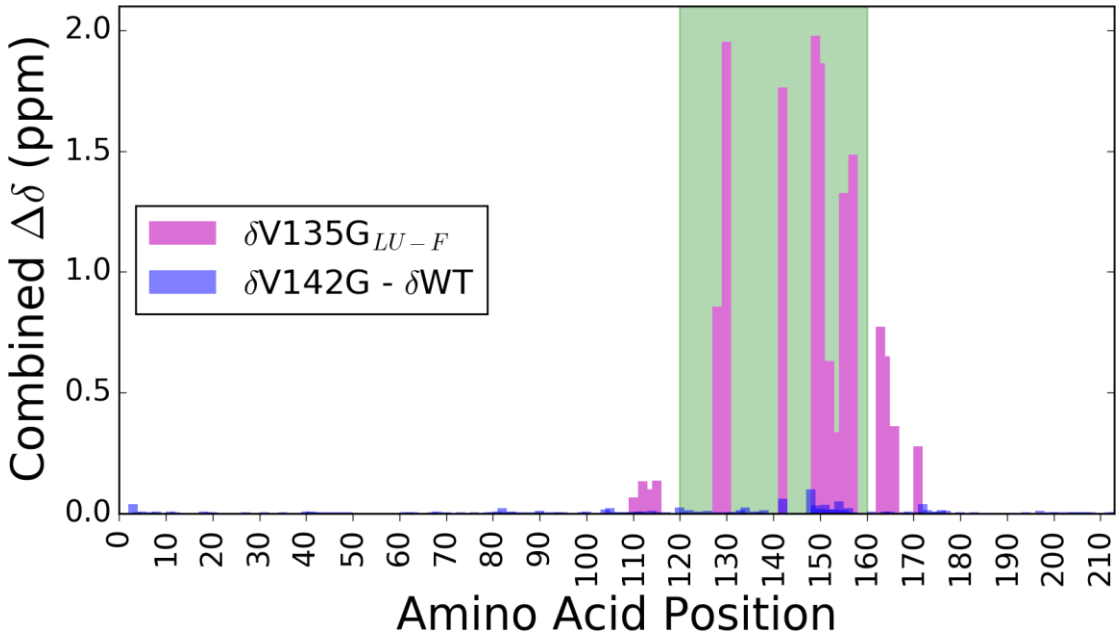

Figure 2.10 V135G Chemical Shift Differences

Combined chemical shift differences between slow exchange pairs in V135G show that exchange is localized to LID domain residues shaded in green. V135G differences are in purple and those between V142G and WT are in blue.

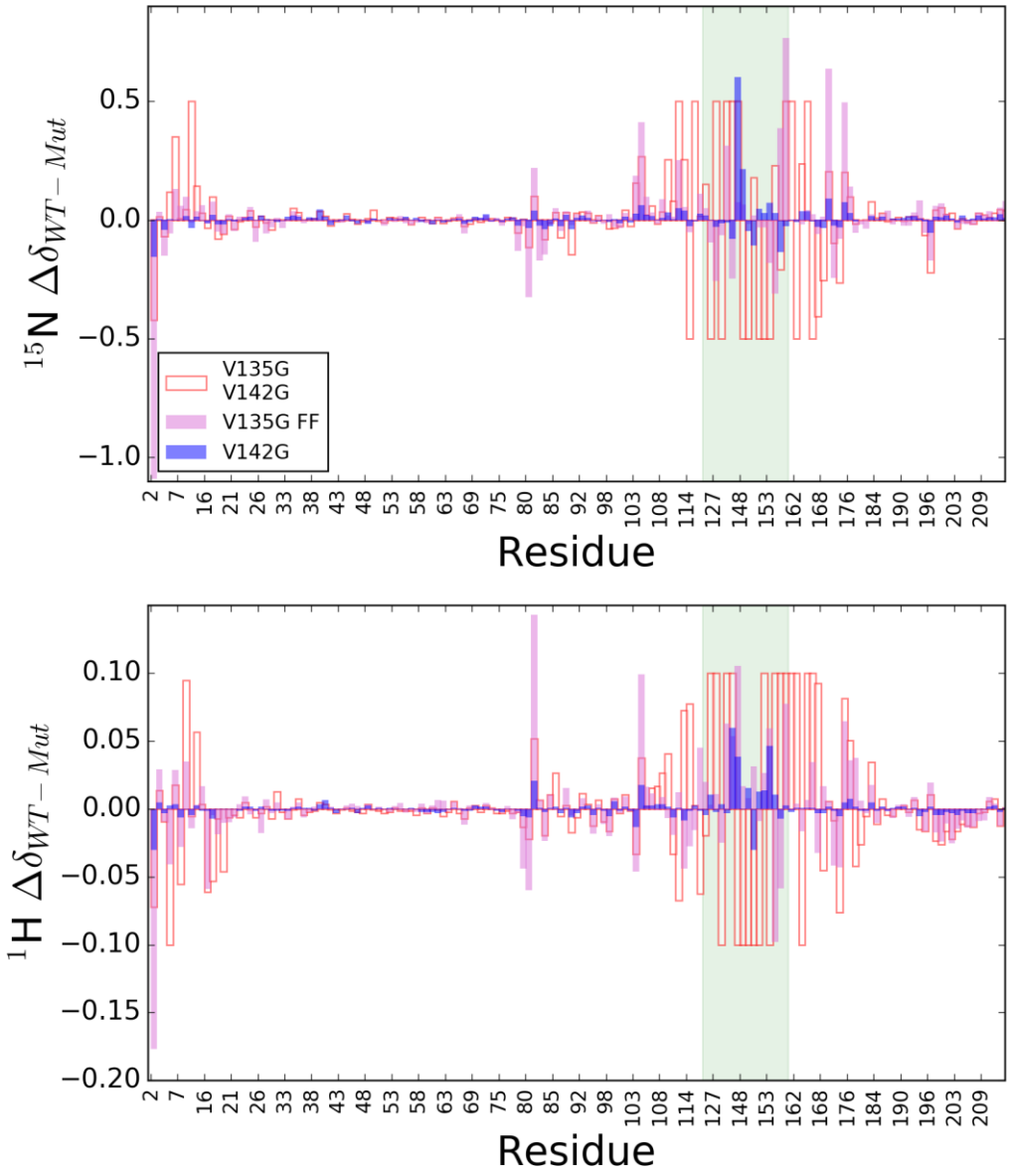

Figure 2.11 Chemical shift differences of folded peaks in V135G

Nitrogen (upper) and proton (lower) chemical shift differences calculated between HSQC spectra collected on WT and all variants at 19 °C show significant changes outside of the LID domain (green) between residues 80-85, 103-108, and 190-201. These regions are all located in the CORE domain. Larger differences were truncated to a magnitude of 0.5 and 0.1 for ^{15}N for and ^{1}H, respectively for the V135G/V142G variant.

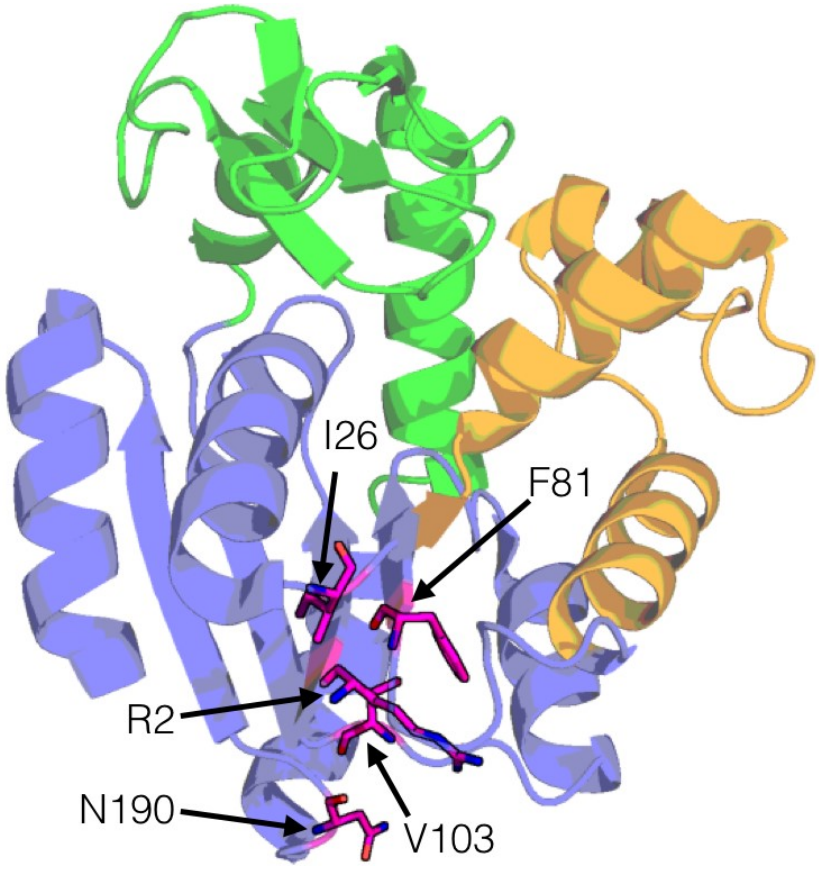

Figure 2.12 Chemical shift differences in CORE residues.

Regions in the CORE domain of AK showing amide chemical shift differences upon entropy-enhancing mutations in the LID domain. Residues shown as sticks in magenta and labeled are those referred to in the text and correspond to x-axis labels in Figure 2.11.

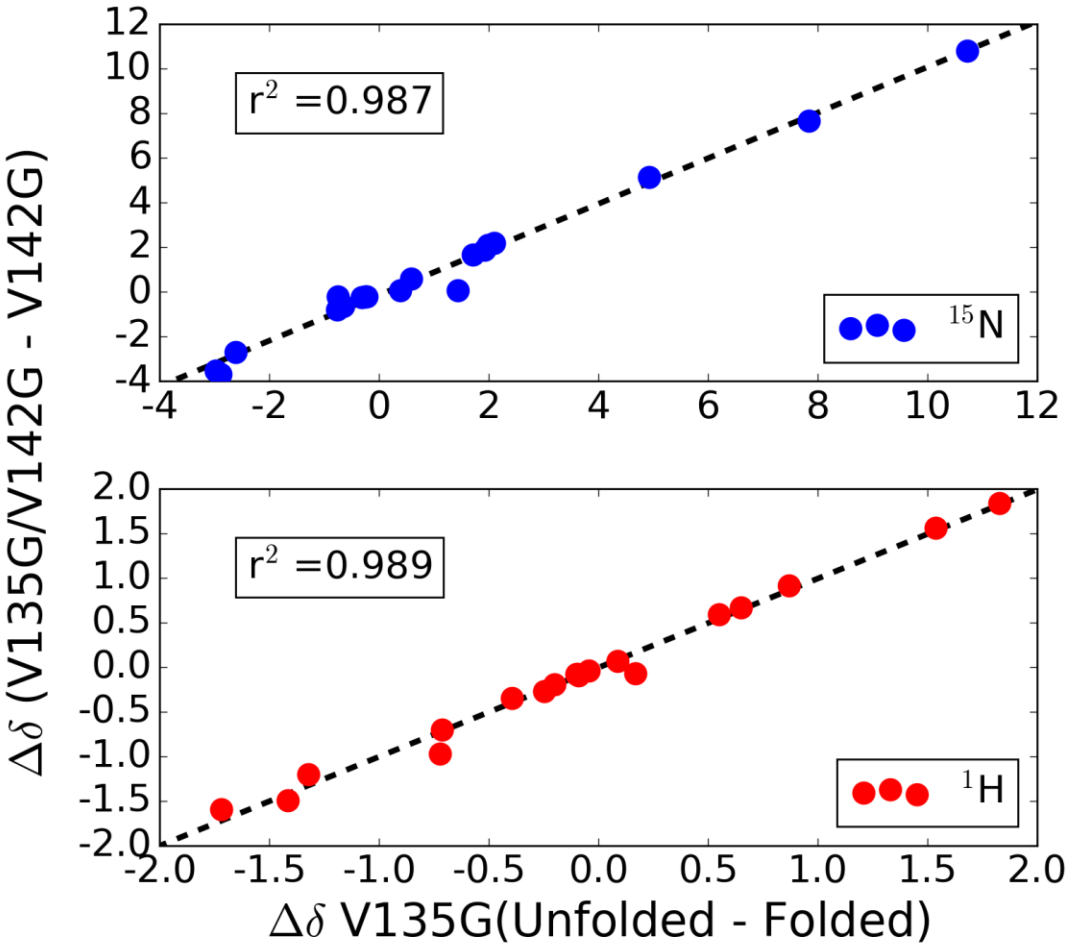

Figure 2.13 Chemical Shift Change Correlation

Correlation of the observed chemical shift differences between slow exchange pairs in V135G (x-axis) and V142G and V135G/V142G (y-axis) demonstrates that all variants access the same local unfolding transition.

Table 2-1 V135G Measured 15N Chemical Shifts, 18.8 T and 19 ℃

	^{15}N $\Delta\delta$ (ppm)	σ (ppm)
6	-0.394	0.19
10	-0.698	0.12
109	-0.269	0.19
111	-0.589	0.14
113	-0.198	0.15
114	0.547	0.26
124	-2.6	0.18
127	-2.738	0.1
129	8.009	0.33
131	-1.244	0.05
141	-2.62	0.16
147	-5.79	0.27
148	5.076	0.15
149	10.818	0.45
151	2.006	0.13
152	2.124	0.06
153	1.677	0.11
154	-0.489	0.14
155	0.659	0.12
156	-2.973	0.34
159	4.608	0.44
161	-3.431	0.19
162	1.906	0.25
164	-0.788	0.2
165	1.708	0.3
172	-2.242	0.15

Chapter 3 – NMR dynamics of local unfolding in V142G and V135G/V142G.

Abstract

Conformational dynamics in proteins can be studied using relaxation dispersion (RD) NMR experiments. These experiments show that LID domain residues in both V142G and V135G/V142G undergo conformational exchange despite being in entirely different conformations. Fit parameters show that exchanging LID residues in both variants have the same chemical shift difference between conformations demonstrating that both variants undergo the same LU transition. The population of the LU state in V142G and V135G/V142G match previous thermodynamic results. Additionally, both variants have similar exchange rate constants on the intermediate to fast NMR timescale. These experiments establish that the LID domain of the V142G and V135G/V142G variants access the locally unfolded and native WT conformations, respectively. Furthermore, this study exemplifies the favorable pairing of entropy-enhancing mutations with NMR experiments measuring protein dynamics, which should be generally applicable.

3.2 Introduction

In addition to structural characterization, NMR spectroscopy has the unique ability to probe conformational dynamics on a per residue basis with minimal or no modifications to the protein. Different NMR experiments can measure a wide range of timescales and obtain information about minor-state populations that are not directly observable.[84, 91, 102] Relaxation dispersion (RD) or CPMG experiments, named after Carr, Purcell, Meiboom, and Gill, are sensitive to dynamics between distinct conformations of proteins on the ms to μs timescale. This process, called conformational exchange, scrambles magnetization, enhances transverse relaxation rates, and reduces peak intensity in HSQC spectra. Spin-

echo refocusing pulses can recover peak intensity by attenuating the effects of conformational exchange and reducing relaxation rates. The profile of the effective transverse relaxation rate, $R_{2,eff}$, as a function of spin-echo pulsing frequency can be fit to analytical models describing the exchange process. Parameters from these fits generally include the populations of the exchanging states, the rate constant of the exchange process, and the chemical shift difference between the states. A more detailed description of this technique is available in the methods section below.

Entropy-enhancing mutations are exquisitely matched with RD experiments because they increase dynamics to enhance lowly-populated states. These experiments provide information about conformations that are not directly observable. Thus, local unfolding dynamics can be explored from the folded and unfolded baselines with the V142G and V135G/V142G variants, respectively. Comparisons of RD fit parameters from multiple variants provide valuable information: chemical shift differences can test for a common structural change, exchange rate constants can inform if folding or unfolding rates increase the LU state, and LU populations from RD and thermodynamics can test for two-state behavior. This chapter demonstrates that both V142G and V135G/V142G undergo the same local unfolding exchange process. These results provide further evidence of the generality of entropy-enhancing mutations and allow for a detailed investigation into the LU conformational transition not otherwise possible.

3.3 Materials & Methods

NMR Relaxation Dispersion

Theory

Relaxation dispersion (RD) is a phenomenon caused by exchange between two distinct conformations of a macromolecule. This process is conceptually depicted in Figure 3.1 with a protein that can access an open and closed conformation. Groups of magnetic moments for these conformations in the transverse plane are depicted by yellow and blue arrows and the green arrow represents the observed,

average chemical shift frequency. During the course of an NMR experiment the two groups of spins are aligned with one another resulting in a maximum average signal. After a time τ one molecule from each group changes conformation, swapping a pair of spins from one group to the other. Single exchange events dephase the spins in each group slightly and reduce the average signal intensity. Transverse magnetization is scrambled if this process continues many times during the T_1 period of an NMR experiment. Experimentally, this effect, called R_{ex}, is added to the intrinsic relaxation processes:

$$R_{2,eff} = R_{2,int} + R_{ex} \tag{2}$$

Conformational exchange results in broader peaks with decreased intensity in 2D spectra like the HSQC. However, the effects of conformational exchange can be mitigated by the application of spin-echo refocusing pulses which are depicted in the lower panel of Figure 3.1. Here only an individual spin from each group is shown as they precess in the transverse plane in the frame of reference of the average shown in green above. The two spins precess in opposite directions over a time τ leading to decreased magnitude of the free induction decay (FID) signal. Next, a 180° pulse is applied, here along the length of the page, such that the x-component of the spins are flipped by 180°. The spins continue to precess in a counter-clockwise direction, due to the static applied magnetic field, and realign after the same time τ. In the context of conformational exchange, the R_{ex} term in (2 can be abated if spin echo pulses are applied fast enough. The magnitude of the R_{ex} term is determined by: the populations of the two states involved, p_A and p_B, the exchange rate constant, k_{ex} (= k_{AB} + k_{BA}), and the chemical shift difference between the two states, $\Delta\omega$, in the following relationship:

$$R_{ex} = \frac{p_A p_B (\Delta\omega)^2}{k_{ex}} \tag{3}$$

In the presence of refocusing pulses the R_{ex} term takes on a more complicated form, which is generally described by the model developed by Carver & Richards in 1972 (CR72)[103]:

$$R_{ex}(\tau) = \frac{k_{ex}}{2} - \frac{\nu_{CPMG}}{2}\cosh^{-1}[D_+\cosh(2\eta_+) - D_-\cos(2\eta_-)] \quad (4)$$

$$D_\pm = \frac{1}{2}\left[\pm 1 + \frac{\psi + 2(\Delta\omega_{AB}^2)}{(\psi^2 + \zeta^2)^{\frac{1}{2}}}\right] \quad (5)$$

$$\eta_\pm = \frac{\left[\pm\psi + (\psi^2 + \zeta^2)^{\frac{1}{2}}\right]^{\frac{1}{2}}}{\sqrt{8}\cdot\nu_{CPMG}} \quad (6)$$

$$\psi = k_{ex}^2 - (\Delta\omega_{AB}^2) \quad (7)$$

$$\zeta = 2\Delta\omega_{AB}k_{ex}(p_A - p_B)\ where\ p_A > p_B \quad (8)$$

Where ν_{CPMG} is the frequency of spin echo pulses applied during the NMR experiment. This model works particularly well for skewed populations in intermediate-to-fast exchange to analyze RD data from the V142G and V135G/V142G variants.

Experiment & Analysis

RD experiments were carried out using the pulse sequence developed by Kay and coworkers using a constant time, during which the spin echo pulses were applied, of 40 and 50 ms for 14.2 and 18.8 T, respectively.[104] Data were collected with relaxation delays of 9.56*, 4.76, 3.16, 1.88, 1.33, 0.92, 0.646*, 0.44, 0.329, 0.28*, and 0.26 ms (26.15, 52.52, 79.11, 132.96, 271.74, 387.17, 568.18, 759.35, 892.86, and 961.54 Hz) at 600 MHz and 9.55*, 4.75, 3.15, 1.87, 1.321, 0.91, 0.636*, 0.43, 0.319, 0.27*, and 0.25 ms (26.18, 52.63, 79.37, 133.69, 189.19, 274.73, 393.26, 581.40, 783.13, 925.93, and 1000.00 Hz) at 800 MHz field strengths. Delays marked with an * were used in duplicate. Batches of TROSY HSQC spectra were collected for each set of delay times in an interleaved manner, to mitigate effects of sample heating, with 16 transients, a 90 ms ^{15}N detection period, and a 2.5 s delay between transients for each relaxation delay at both 14.2 and 18.8 T field strengths at 19°C.

Interleaved data were separated using in-house scripts, FIDs were processed with nmrPipe, and assignments were transferred using the CCPNMR Analysis software.{Skinner, 2015 #1008;Skinner, 2016 #1007} Peak lists were exported and the CR72 model fit to the data using nonlinear-least-squares with the stand-alone python software package *relax*.[105] Transverse relaxation rates were calculated within *relax* using the expression:

$$R_{2,eff} = \frac{\ln\left(\frac{I_\nu}{I_0}\right)}{CT} \tag{9}$$

where I_0 is the intensity of the reference TROSY HSQC spectrum, I_ν the intensity of the spectra collected with the CPMG spin echo pulse train applied, and CT the constant time period during the CPMG pulse train. The two-point experiment to estimate the magnitude of the exchange contribution to the relaxation rate was determined by taking the difference between the $R_{2,eff}$ calculated for highest and lowest pulsing frequency. High pulsing frequency abolishes the R_{ex} term leaving only $R_{2,int}$. Subtracting the two $R_{2,eff}$ terms leaves R_{ex} since $R_{2,int}$ is the same at both pulsing frequencies.

$$R_{ex} = R_{2,eff}^{\nu,max} - R_{2,eff}^{\nu,min} \tag{10}$$

The relax software program was used to analyze full RD curves from both field strengths using the CR72 model described above. First, a grid search was performed using 21 points over defined ranges for each fit parameter: R2 = 5-20 s^{-1}; $\Delta\delta$ = 0-14 ppm; p_A = 0.5-1.0; k_{ex} = 1-10000. This coarse-grained search protected against selection of a local minimum in parameter space. Local minima were found using a Levenberg-Marquardt, nonlinear-least-squares fit, starting with the optimal grid-search parameters. Parameter error estimation was performed using Monte Carlo permutation tests. In this process $R_{2,eff}$ values are calculated for the same measured pulsing frequencies with the CR72 model with normally distributed error added to the value and fit using the same model. This process was repeated 500 times. Data, fits, and parameters were plotted using ipython and matplotlib.

3.4 Results & Discussion

V142G & V135G/V142G LID Dynamics

Conformational fluctuations can be present in a protein but not observable in HSQC spectra. Disappearance of LID peaks at high temperatures in V142G implies that these residues are dynamic and structural data from V135G and V135G/V142G suggests that the V142G LID accesses an unfolded conformation. Relaxation dispersion (RD) experiments, as described in the methods section above, were performed on V142G to investigate dynamic LID unfolding. These experiments can indirectly detect lowly-populated states like the LU conformation that are unobservable in HSQC spectra through observations of a single, observed HSQC peak in a different conformation. A two-point experiment, described by equation (10) and shown in Figure 3.2 A, can quickly probe for dynamics and estimate the magnitude of the exchange component of transverse relaxation, R_{ex}. Figure 3.2 B shows this estimate along the sequence of AK with the WT in gray and V142G in blue. Although V142G is structurally the same as WT at 19 °C (see Figure 2.2), large R_{ex} values are present in the LID domain, shaded in green. These results demonstrate that the V142G LID is undergoing conformational exchange while the WT protein does not. Interestingly, glycine 10, in the CORE domain, is also dynamic. This residue is positioned in the canonical ATP binding p-loop in the CORE domain that is roughly in between the two points of insertion of the LID as shown in **Error! Reference source not found.**. The p-loop is vital to the binding of substrate to AK and conformational fluctuations within this region may cause the decreased binding affinity in entropy-enhancing variants.

Full analysis of RD data provides structural, kinetic, and thermodynamic information about the conformations in exchange. This information is extracted from fits to RD curves, as shown in Figure 3.2 A, using previously developed numeric and analytical models. One of the first and most general models was developed by Carver and Richards in 1972, referred to here as the CR72 model, and is described by equations (4)-(7). Fit parameters from the CR72 model include: the major state population, p_A, the

exchange rate constant, k_{ex}, and the chemical shift difference between the two states, $\Delta\omega$. Multiple field strengths are vital for accurate determination of parameters, as the chemical shift difference is field dependent. CR72 was chosen from a set of plausible models using the Aikake information criterion. Fits to V142G RD curves from exchanging residues with well-resolved peaks and an estimated R_{ex} of at least $5s^{-1}$ are shown in Figure 3.3 and Figure 3.4 and described in Table 3-1. Most of these curves plateau to intrinsic R_2 values at high pulsing frequency; however, residues that do not show a plateau include 129, 136-138, 147-149, and 158-159. This behavior could be due to faster exchange rates but is more likely a result of larger ^{15}N chemical shift differences as measured in V135G (Table 2-1). The model fits poorly to some residues at low pulsing frequency for both 600 (blue) and 800 MHz data (red). It is not clear if this is due to experimental error or limitations of the model. Systematically high $R_{2,eff}$ values were also observed at 925.93 and 1000 Hz pulsing frequencies in 800 MHz data and were omitted from all analyses in this study. The CR72 fits recapitulate the main features of the data.

V142G RD fit parameters are consistent with LID unfolding. Reduced χ^2 values, shown in Figure 3.4 A, demonstrate that most residues undergoing exchange are described well by the CR72 model. Populations and exchange rate constants should cluster around a common value for multiple residues if the exchange process is two-state and cooperative. The population of the folded state in Figure 3.4 B, p_A for this analysis, clusters near the median value of 0.89, which is consistent with previous thermodynamic data.[81] A handful of residues clustered around 0.5, which may indicate a limitation of the model, as this is the minimum allowed value in the data analysis. Exchange rate constants are quite variable ranging from 25 to 1100 s^{-1} with a median value of 321 s^{-1}, as seen in Figure 3.4 C. The variability of both exchange rates and populations prevented a cluster analysis of the data in which these parameters are shared amongst all exchanging residues. Notably, the largest deviations from the median and errors in k_{ex} values occur for residues with large $\Delta\delta_{fit}$ values, which could be due to poor fitting to the data, actual differences in values, or evidence for multi-state exchange. To this last point, residues with large $\Delta\delta$ values will have larger R_{ex} values and a greater signal from an additional exchange process. Figure 3.4 D shows nitrogen chemical shift differences, which, unlike the rates and populations, should not cluster about a common value. Many

of these differences are quite large, which is consistent with unfolding, and match directly measured values as discussed below. The fit parameters described here are consistent with the LID domain in V142G locally unfolding on an intermediate-to-fast timescale at a population of ~10% at 19°C.

RD experiments were also performed investigate dynamics in V135G/V142G. Although this variant has an unfolded LID, it undergoes conformational exchange in the exact same regions as V142G as seen in Figure 3.2. Fewer residues in V135G/V142G were analyzed owing to spectral overlap but many have well-resolved peaks and an estimated R_{ex} of at least $5s^{-1}$as seen from the curves and fits to the CR72 model in Figure 3.5. V135G/V142G fit parameters describe an intermediate-to-fast exchange process in the LID domain with large chemical shift differences and a folded population of only 3%. Reduced χ^2 values in Figure 3.6 A show that fits to V135G/V142G are comparable to those to V142G data. The major-state population in Figure 3.6 B, p_A, describes the *unfolded* population for the V135G/V142G variant and clusters well about a value of 0.97. This small folded population is consistent with an undetectable folded state population by thermodynamics. Exchange rate constants in Figure 3.6 C have a median value of 394 s^{-1}. This rate constant is very close to that measured in V142G, which means, assuming a two-state process, that folding and unfolding rates decrease and increase, respectively, to change populations in V135G/V142G. Similarly to V142G, rates are highly variable and k_{ex} errors correspond to residues with large $\Delta\delta_{fit}$ values. ^{15}N chemical shift differences are large and similar to values from V142G data, implying that the unfolded V135G/V142G LID is accessing a folded conformation.

A similar transition in V142G & V135G/V142G

Parameters from V142G and V135G/V142G are consistent with a common LID unfolding process. Although different conformations are observed in these variants, ^{15}N chemical shift differences are highly correlated between the folded, V142G and the unfolded, V135G/V142G variants as seen in Figure 3.7. Two outliers, K136 and F137, present for V135G/V142G, have much lower values and likely differ owing to local effects of the mutation since these residues are adjacent to V135. The $\Delta\delta$ fit parameters for both V142G and V135G/V142G can be compared to measured $\Delta\delta$ values between slow-exchange pairs in

V135G. This comparison is shown in Figure 3.8 with chemical shift differences in blue and red for V142G and V135G/V142G, respectively, linear fits as dashed lines, and a grey dashed line along the diagonal. Fits to each variant are close to the ideal diagonal line demonstrating excellent agreement between measurements and fit parameters over a large range of chemical shift differences. However, shift differences are systematically lower and higher than those observed in V135G for V142G and V135G/V142G, respectively. This result, while subtle, may be evidence of multi-state exchange as described in Chapter 4 below.

Conclusions

The above results demonstrate that the LID domains of V142G and V135G/V142G are dynamic and access the same conformations. Both variants show dynamics on intermediate-to-fast timescales, have populations that match previous thermodynamic measurements, and have the same chemical shift differences. This last point is remarkable because the observed LID residues are in entirely different conformations. Furthermore, the fact that both of these variants agree with directly observed chemical shift changes from V135G further strengthens the fact that all variants are undergoing the same process. Lastly, this study shows that entropy-enhancing mutations are a useful tool for dynamics measurements. In fact, experiments were performed at 19 °C because the V142G variant matches the LU population of WT at 37°C but has measurable dynamics. Coupled with minor perturbations to the ground state, entropy-enhancing mutations should prove useful in many cases.

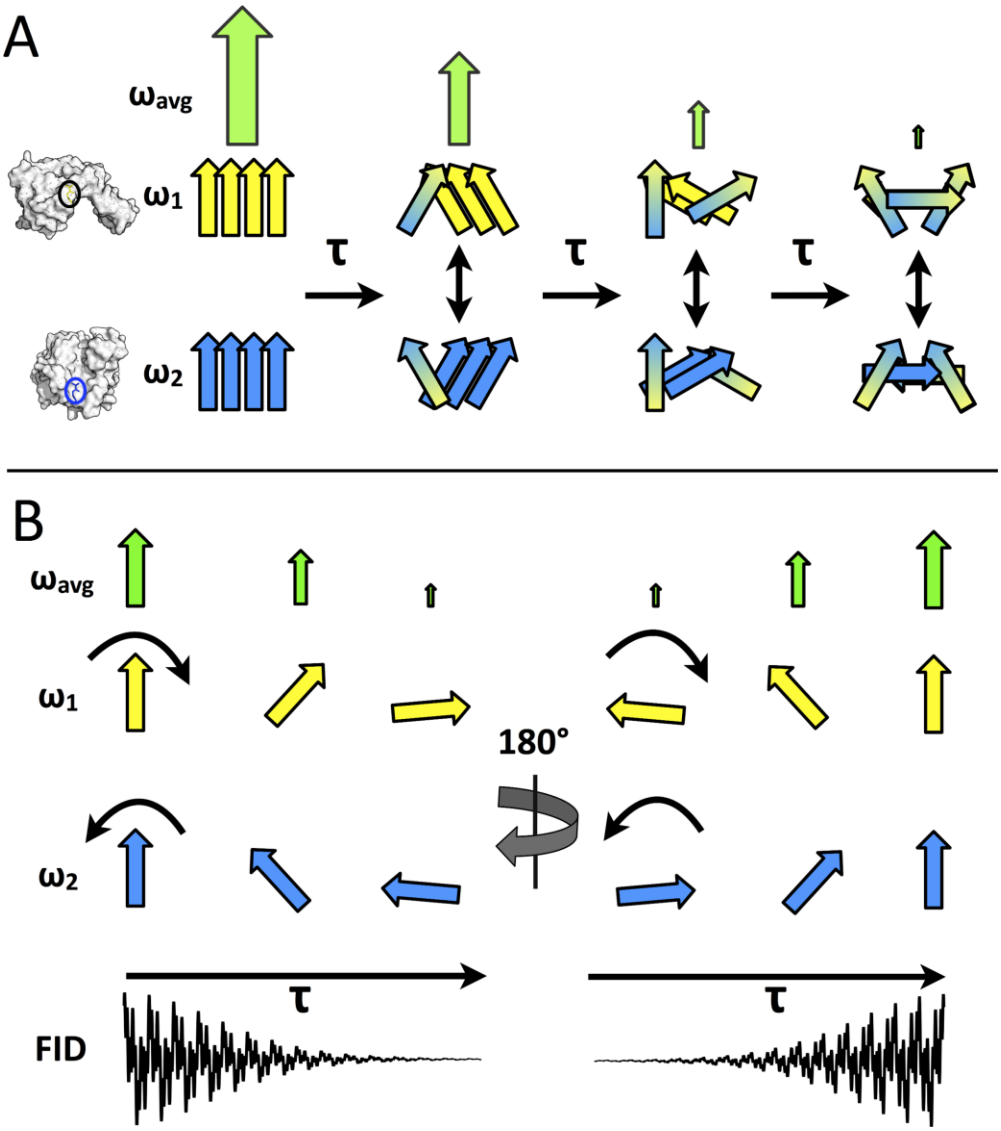

Figure 3.1 CPMG and Spin Echo Concept

A. Conformational exchange scrambles transverse magnetization. The diagram follows two groups of conformational states of the same spin (yellow and blue arrows) in the rotating frame of reference of the average chemical shift frequency (green arrows). The two states precess at slightly different frequencies and as individual proteins change conformations after a time τ the groups of spins dephase giving rise to the R_{ex} contribution to the transverse relaxation rate. B. Spin echo pulses ameliorate dephasing due to conformational exchange. A 180° pulse (along y in this case) can be applied to the system to refocus the spins of the two conformations in A prior to conformational exchange. After the pulse the spins are flipped in the transverse plane but precess in the same direction as before the pulse.

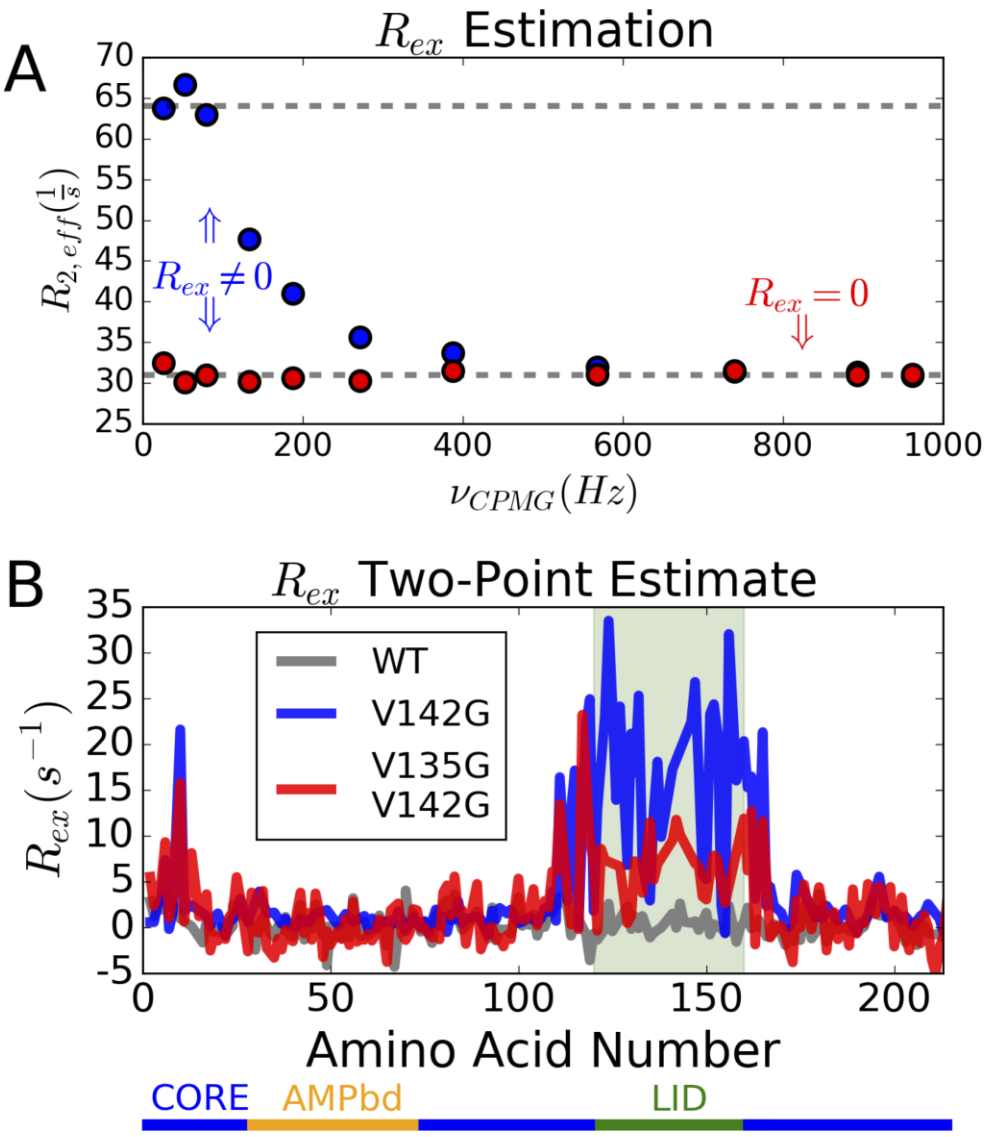

Figure 3.2 R_{ex} Two-Point Estimate

A. RD curves from V142G showing evidence of exchange as changes in $R_{2,eff}$ (blue, E152) and no exchange (red, T155). The two-point estimate of the R_{ex} contribution is calculated by subtracting the highest ν_{CPMG} data point from the lowest. B. The two-point R_{ex} estimate shows exchange in the LID domain (green) and N-terminus in V142G (blue) and V135G/V142G (red) but not in the WT (grey).

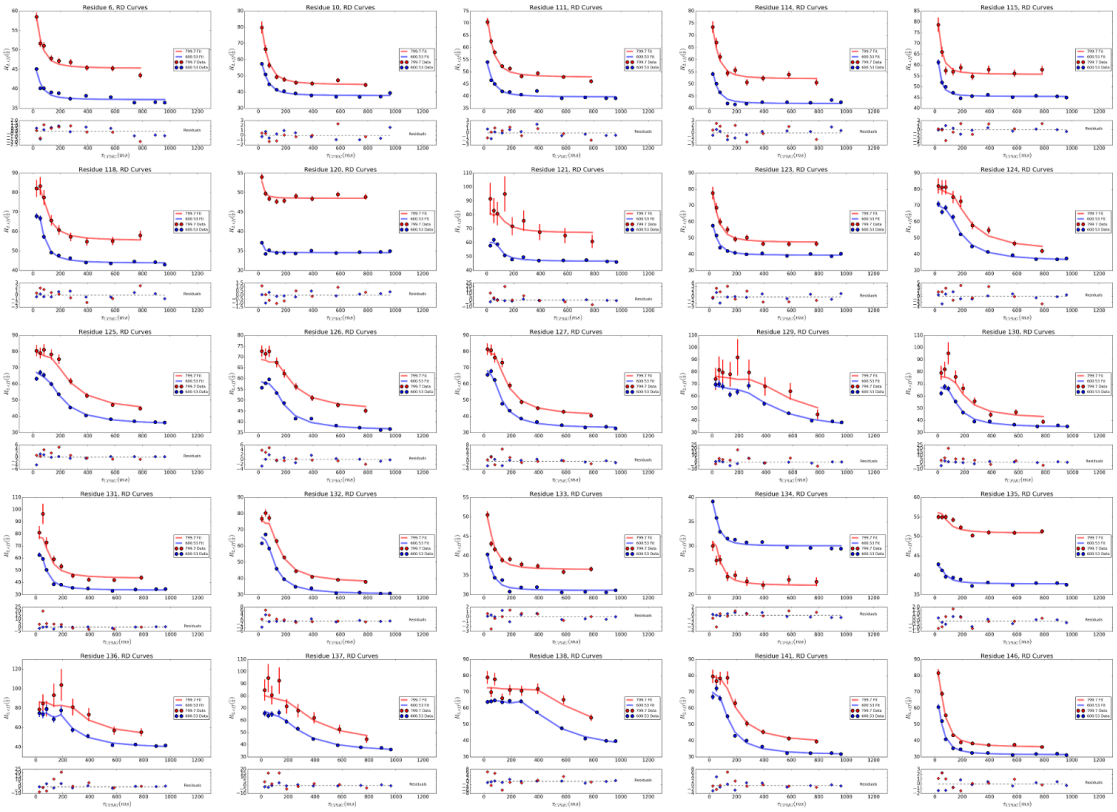

Figure 3.3 V142G RD Curves

RD curves for residues in V142G that show appreciable exchange and have well-resolved peaks and estimated R_{ex} values of at least $5s^{-1}$. Data are shown as dots, with errors propagated from the variance in peak intensity, and fits as lines. Data were collected at 14.2 and 18.8 T in blue and red, respectively. Only these residues and those in Figure 3.44 were used for further analysis.

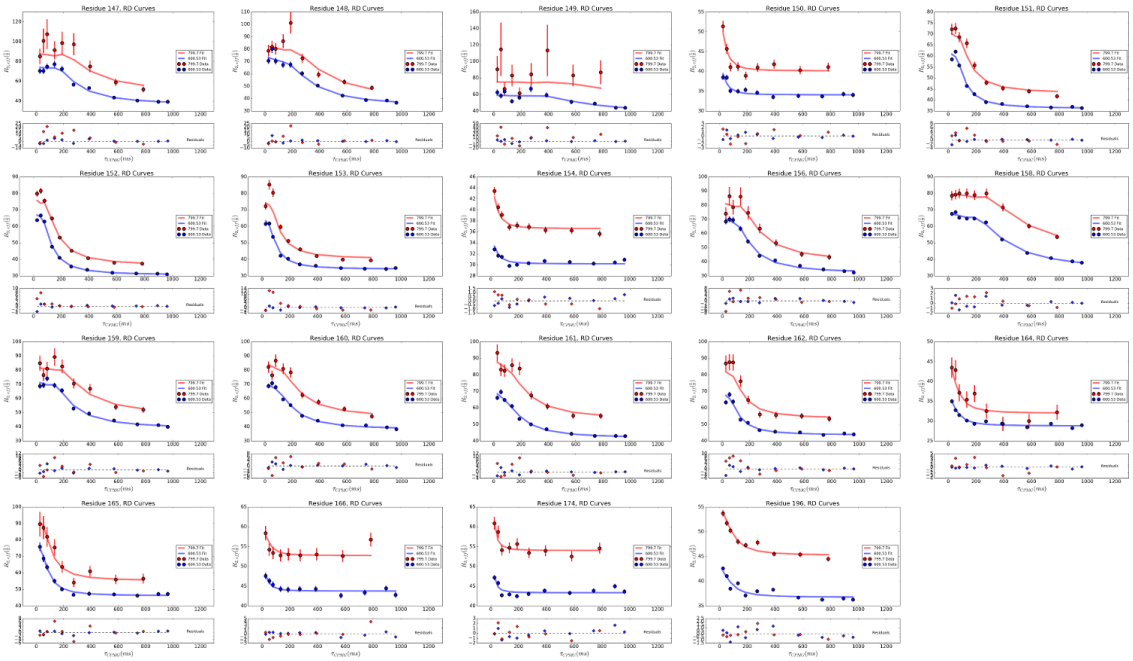

Figure 3.4 Contd. V142G RD Curves

V142G RD curves for residues 147 through 196.

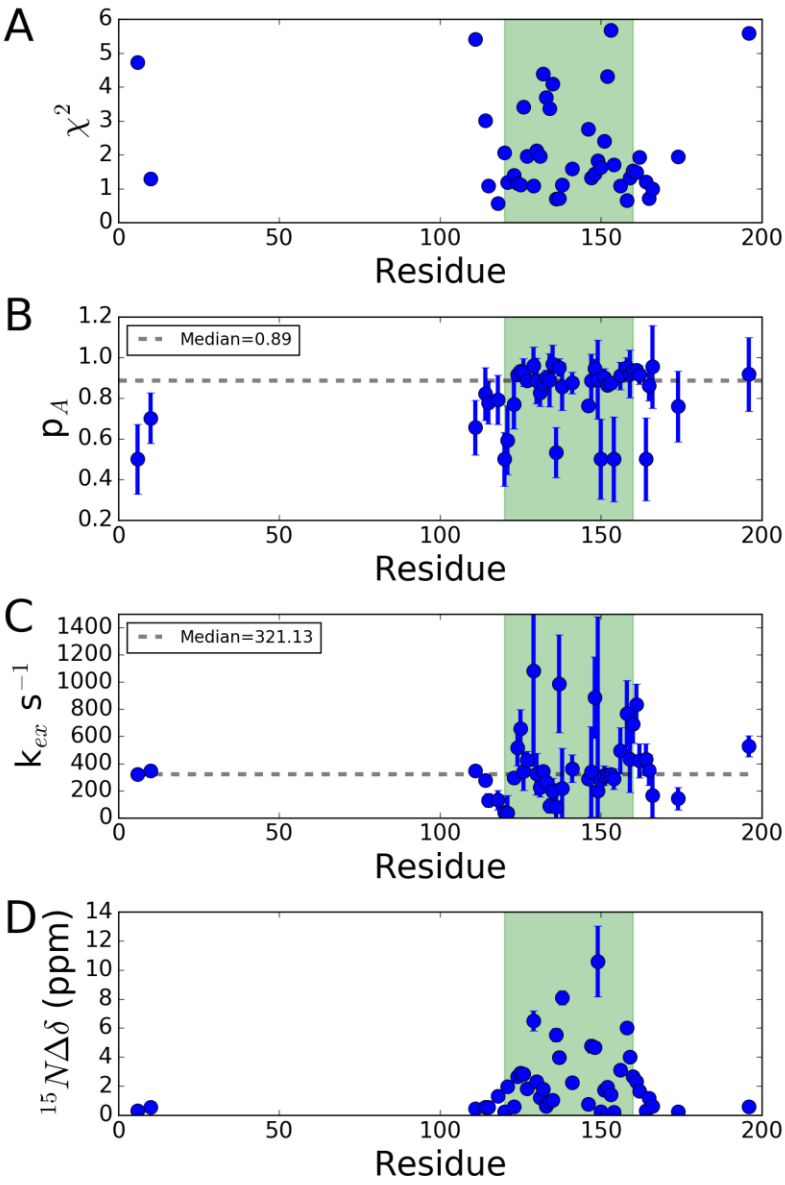

Figure 3.4 V142G CR72 Fit Parameters

A. Reduced χ^2 values show that the CR72 model fits well to V142G data. B. Major state populations, the folded conformation here, are consistent with previous thermodynamic data and have a median value for LID residues of 0.89, shown as a grey dashed line. C. Kinetic rate constants for the exchange process have a median value (dashed line) of 321 s^{-1} for LID residues but are highly variable. D. ^{15}N chemical shift differences are largest within the LID domain. Errors are derived from Monte Carlo permutation tests in all cases and the green region highlights LID residues.

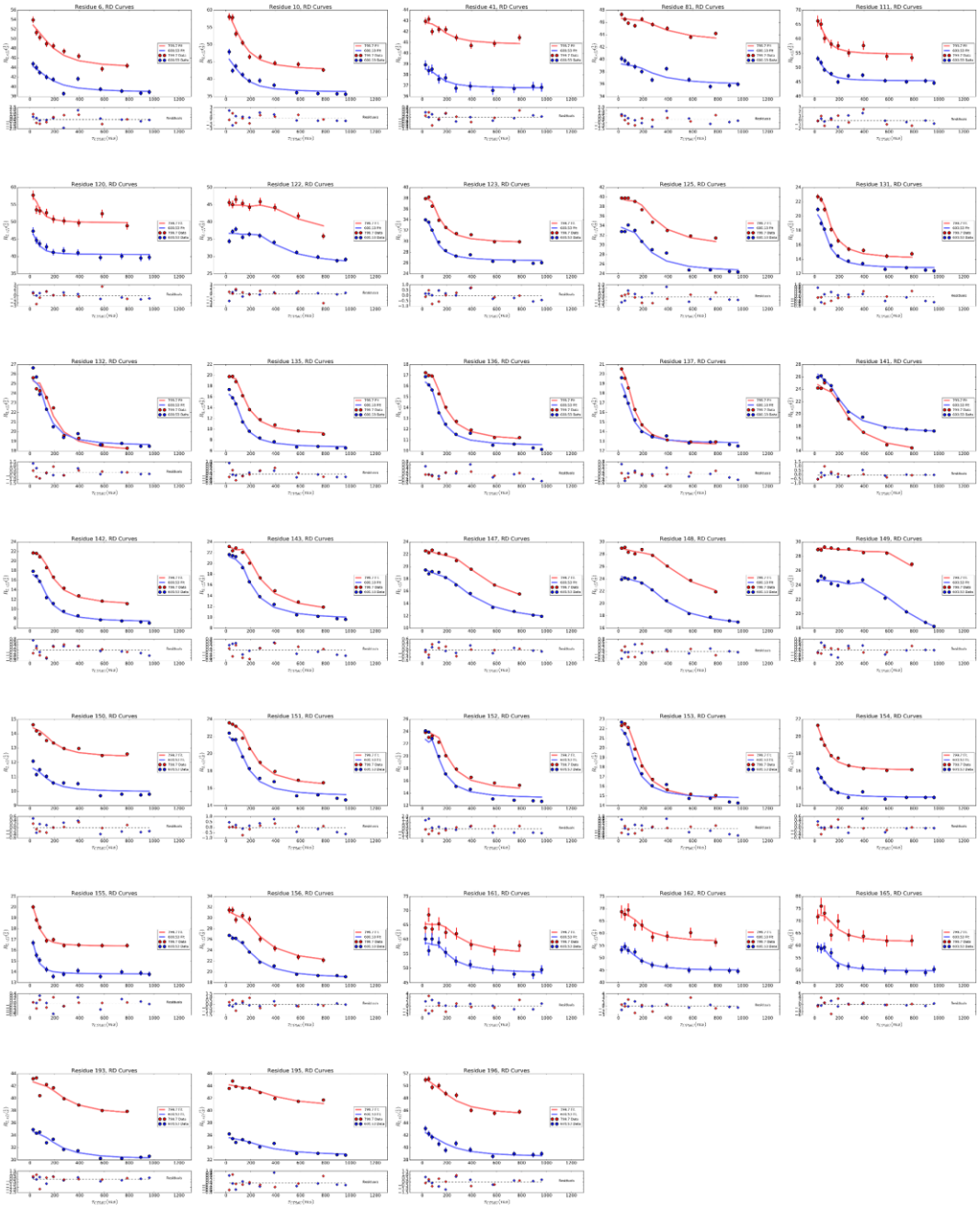

Figure 3.5 V135G/V142G RD Curves

RD curves for V135G/ V142G with appreciable exchange and have well-resolved peaks and estimated R_{ex} values of at least $5s^{-1}$. Dots represent data and fits as lines. Errors were propagated from the variance in peak intensity. 14.2 and 18.8 T data are in blue and red, respectively. Only residues shown here were used for further analysis.

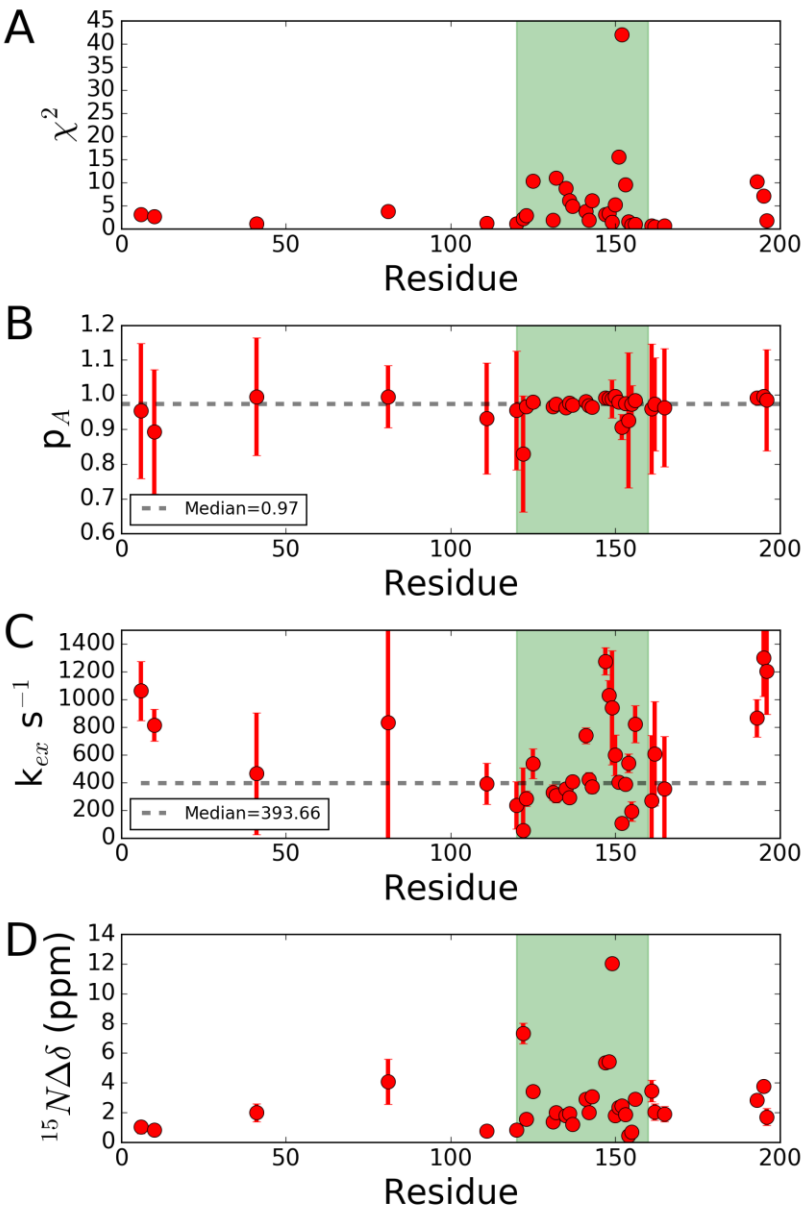

Figure 3.6 V135G/V142G RD Parameters

A. Reduced χ^2 values show that the CR72 model fits well to the data. B. Major state populations, representing the unfolded conformation, have a median value (dashed line) of 0.97 for LID residues. B. Kinetic rate constants for the exchange process have a median value (dashed line) of 394 s^{-1} for LID residues but are highly variable. C. ^{15}N chemical shift differences are well determined and largest within the LID domain. In all cases errors are derived from Monte Carlo permutation tests and the green region highlights LID residues.

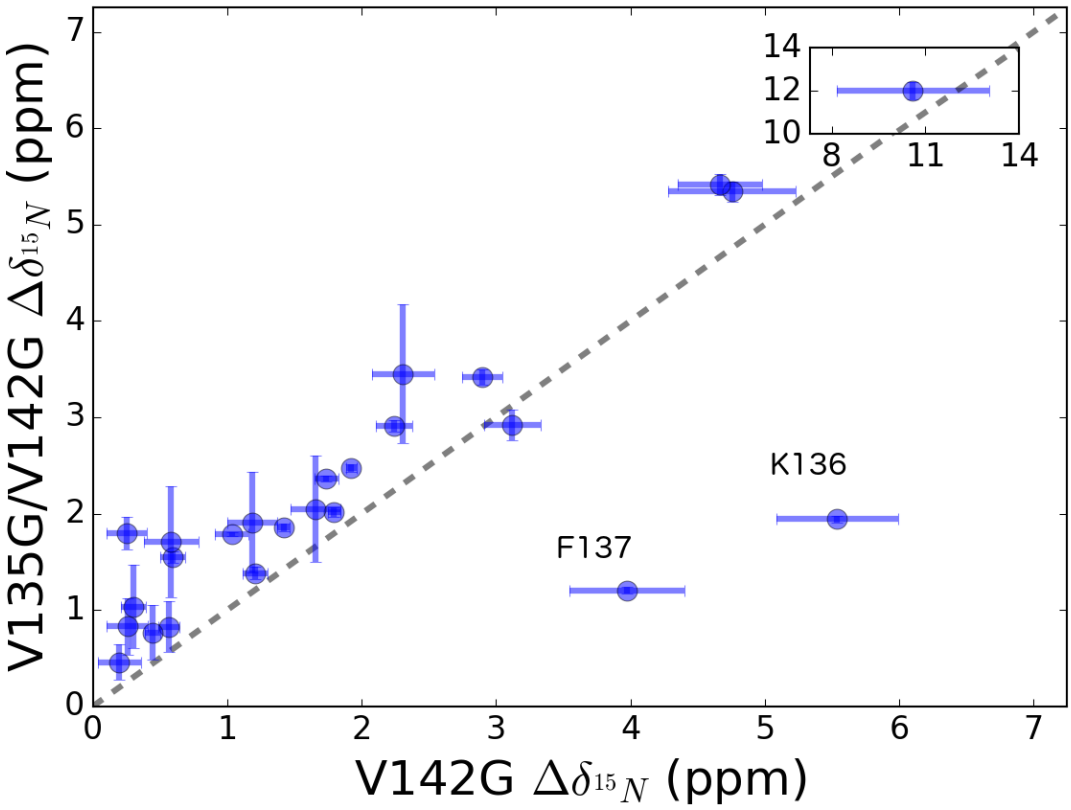

Figure 3.7 Δδ Fit Comparison

^{15}N chemical shift differences from RD fits correlate well for common residues between V142G and V135G/V142G. This provides strong evidence that both variants are engaged in the same local unfolding process. However, V142G has systematically smaller values than V135G/V142G, which could be due to multi-state exchange. Furthermore, K136 and F137 are outliers in the V135G/V142G variant based on measured values in V135G, likely because of local effects of the V135G mutation.

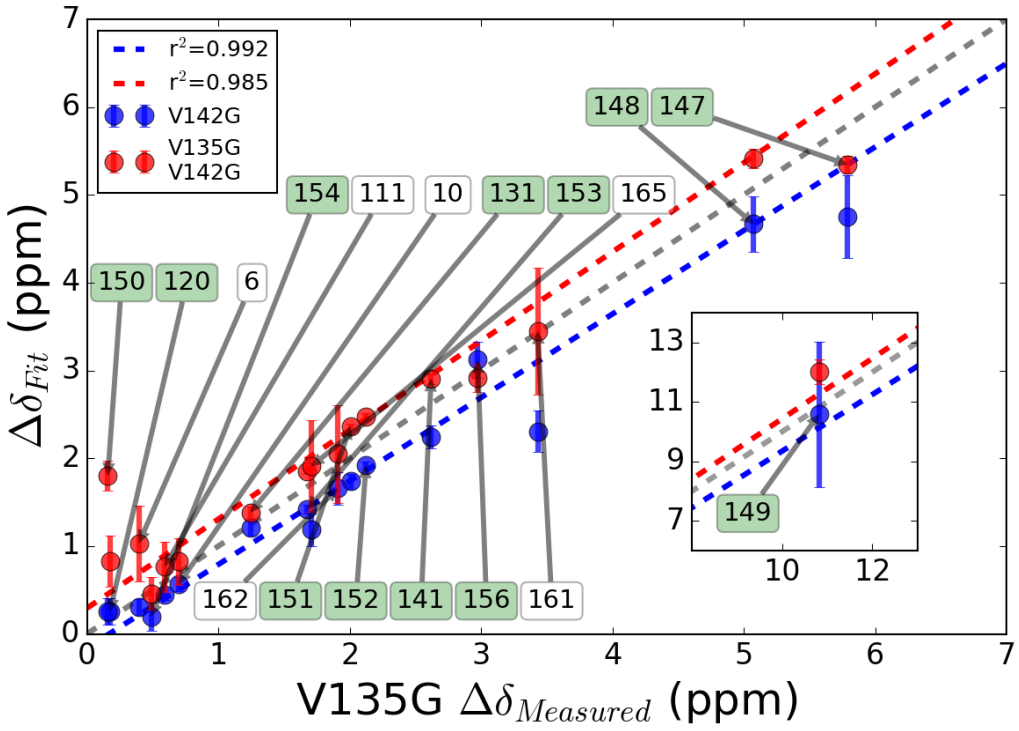

Figure 3.8 RD Δδ Comparison

^{15}N chemical shift differences between slow exchange pairs in V135G correlate with fit parameters from both V142G (blue) and V135G/V142G (red) establishing that all variants access the same dynamic process despite only one observable population in V142G and V135G/V142G. Data points are labeled with residue numbers in boxes and labels for LID residues, from 120-160, are shaded in green. Slightly larger and smaller values for V135G/V142G and V142G, respectively could be due to experimental error or evidence of multi-state exchange.

Table 3-1 V142G CR72 CPMG Fit Parameters

Parameters corresponding to Figure 3.4 from fits to V142G CPMG curves presented in Figures Figure 3.3.

	χ^2	p_A	p_A error	$k_{ex}(s^{-1})$	k_{ex} error	$\Delta\delta$ (ppm)	$\Delta\delta$ error
6	4.73	0.5	0.17	322.04	39.97	0.3	0.09
10	1.29	0.7	0.12	344.45	39.6	0.57	0.08
111	5.41	0.66	0.13	346.86	28.61	0.44	0.06
114	3.01	0.82	0.13	277.15	44.81	0.54	0.11
115	1.09	0.78	0.1	129.02	49.92	0.56	0.1
118	0.57	0.79	0.12	130.47	72.84	1.31	0.15
120	2.07	0.5	0.13	37.45	50.16	0.26	0.16
121	1.2	0.59	0.17	34.39	129.5	1.97	0.29
123	1.41	0.77	0.12	294.82	45.73	0.59	0.09
124	1.16	0.92	0.04	515.34	133.47	2.67	0.17
125	1.12	0.93	0.01	658.08	138.77	2.9	0.15
126	3.41	0.93	0.07	343.18	140.85	2.83	0.15
127	1.96	0.89	0.01	425.67	63.44	1.83	0.08
129	1.09	0.96	0.09	1083.75	703.55	6.51	0.69
130	2.12	0.89	0.11	323.8	153.78	2.33	0.24
131	1.96	0.83	0.07	222.51	68.85	1.21	0.09
132	4.39	0.87	0.01	340.93	41.6	1.8	0.05
133	3.7	0.9	0.04	263.65	41.81	0.62	0.09
134	3.36	0.89	0.13	87.2	53.69	0.93	0.09
135	4.09	0.97	0.09	194.48	96.71	1.03	0.13
136	0.7	0.53	0.12	79.06	176.37	5.54	0.45
137	0.71	0.95	0.04	988.54	362.45	3.98	0.43
138	1.12	0.86	0.12	216.8	295.62	8.09	0.48
141	1.6	0.88	0.05	362.72	101.09	2.24	0.13
146	2.76	0.76	0.02	285.73	20.45	0.77	0.03
147	1.33	0.89	0.13	340.04	333.9	4.76	0.47
148	1.44	0.95	0.05	887.24	299.64	4.67	0.32
149	1.83	0.89	0.19	197.13	1287.57	10.59	2.44
150	1.63	0.5	0.2	281.25	77.61	0.25	0.15
151	2.4	0.9	0.04	310.04	72.66	1.74	0.09
152	4.32	0.86	0.01	318.15	34.01	1.92	0.04
153	5.68	0.87	0.01	321.13	36.95	1.42	0.05
154	1.7	0.5	0.21	285.46	74.67	0.2	0.16
156	1.09	0.91	0.07	493.67	171.4	3.12	0.21
158	0.66	0.95	0.05	769.25	244.34	6.01	0.23
159	1.33	0.92	0.12	431.41	245.52	4.0	0.36
160	1.53	0.93	0.02	691.2	141.72	2.65	0.17
161	1.49	0.94	0.01	834.83	150.92	2.31	0.23
162	1.93	0.92	0.05	420.31	126.72	1.66	0.18
164	1.2	0.5	0.2	430.88	114.67	0.31	0.24
165	0.72	0.86	0.08	350.79	113.47	1.19	0.19
166	1.0	0.96	0.2	166.3	165.74	0.61	0.31
174	1.95	0.76	0.17	141.56	83.61	0.25	0.2
196	5.6	0.92	0.18	527.71	77.82	0.58	0.2
mean	2.16	0.82	0.1	375.15	157.88	2.27	0.25
std	1.42	0.15	0.06	244.1	213.08	2.28	0.37
min	0.57	0.5	0.01	34.39	20.45	0.2	0.03
max	5.68	0.97	0.21	1083.75	1287.57	10.59	2.44

Table 3-2 V135G/V142G CR72 CPMG Fit Parameters

Parameters plotted in Figure 3.6 from fits to CPMG curves collected on the V135G/V142G variant shown in Figure 3.5.

	χ^2	p_A	p_A error	$k_{ex}(s^{-1})$	k_{ex} error	$\Delta\delta$ (ppm)	$\Delta\delta$ error
6	3.11	0.95	0.2	1059.4	212.36	1.03	0.43
10	2.68	0.89	0.18	813.71	114.45	0.83	0.27
41	1.13	0.99	0.17	463.97	438.58	1.99	0.6
81	3.85	0.99	0.09	830.13	1151.78	4.08	1.53
111	1.26	0.93	0.16	390.48	149.45	0.76	0.28
120	1.16	0.95	0.17	236.33	170.57	0.83	0.29
122	2.25	0.83	0.17	54.3	449.49	7.33	0.7
123	2.92	0.97	0.01	283.78	54.24	1.55	0.06
125	10.41	0.98	0.01	535.04	107.58	3.41	0.08
131	1.94	0.97	0.0	327.57	54.63	1.38	0.07
132	11.1	0.97	0.01	306.89	51.84	2.01	0.05
135	8.87	0.96	0.0	353.9	19.82	1.79	0.02
136	6.18	0.98	0.0	289.68	36.23	1.94	0.03
137	4.98	0.97	0.0	404.37	26.12	1.2	0.04
141	3.8	0.98	0.0	737.78	57.11	2.91	0.06
142	1.9	0.97	0.0	419.75	51.15	2.0	0.06
143	6.2	0.96	0.0	370.32	51.39	3.07	0.04
147	3.15	0.99	0.0	1273.94	99.28	5.34	0.1
148	3.42	0.99	0.0	1029.2	106.35	5.41	0.11
149	1.47	0.99	0.05	939.66	411.36	12.02	0.4
150	5.24	0.99	0.02	596.8	145.05	1.8	0.17
151	15.56	0.98	0.0	400.53	32.72	2.37	0.03
152	42.09	0.91	0.04	104.58	38.15	2.47	0.04
153	9.66	0.97	0.0	386.8	33.12	1.86	0.04
154	1.57	0.93	0.2	537.9	67.46	0.46	0.18
155	0.83	0.97	0.05	190.16	69.39	0.7	0.1
156	1.02	0.98	0.0	821.98	133.28	2.92	0.16
161	0.67	0.96	0.19	269.8	469.12	3.45	0.72
162	0.5	0.97	0.13	606.83	377.08	2.05	0.55
165	0.69	0.96	0.17	354.52	377.92	1.91	0.52
193	10.24	0.99	0.0	863.74	134.76	2.82	0.15
195	7.15	1.0	0.0	1298.62	279.91	3.76	0.29
196	1.87	0.98	0.15	1200.78	310.8	1.71	0.57
mean	5.42	0.96	0.07	568.28	190.38	2.7	0.27
std	7.58	0.03	0.08	341.23	224.03	2.25	0.31
min	0.5	0.83	0.0	54.3	19.82	0.46	0.02
max	42.09	1.0	0.2	1298.62	1151.78	12.02	1.53

Chapter 4 – Conformational dynamics in V135G and multi-state exchange.

4.1 Abstract

In this chapter measurements of dynamics in the V135G variant are presented and the effects of entropy-enhancing mutations are analyzed in the context of two- and three-state exchange. The V135G variant is in slow exchange conformations and both ZZ exchange and RD experiments measure exchange kinetics on the order of 20 s^{-1}. However, 2-state models are unable to fit ZZ exchange data and rates measured from these two experiments are different. Additionally, a van't Hoff analysis was performed using NMR data and the results do not agree with previous calorimetric values indicating that the process is not 2-state. Additional circumstantial evidence from V135G is consistent with two parallel unfolding processes with similar structural endpoints, one slow and the other intermediate-to-fast on the NMR timescale. Evidence for slow exchange in the V142G and V135G/V142G variants is presented that supports this multi-state exchange model. The effects of mutations in the context of a 2-state model are considered and strategies for testing multi-state exchange are discussed.

4.2 Introduction

The V135G variant shows clear evidence of slow exchange between the native and LU conformations. In addition to verifying exchange, ZZ exchange experiments can be used to quantify conformational dynamics.[89, 106-107] Conformations are allowed to exchange during the experiment generating cross peak intensities that are dependent on the mixing time during which exchange occurs. Rates of exchange can be obtained by fitting analytical expressions to ZZ exchange data or fitting a line to the linear portion of cross peak intensities at short mixing times. To obtain rates from ZZ exchange

experiments, self and cross peaks must be resolved in HSQC spectra, which is difficult with AK because the spectrum is crowded due to the size of the protein and the fact that unfolded peaks are not well dispersed. Although peaks for multiple residues are well resolved, two-state models do not fit the ZZ exchange data. On the other hand, relaxation dispersion experiments can measure dynamics for residues in slow exchange and only require a single peak for quantitative analysis. The setup of RD experiments and analysis of the data is the same as that described for V142G and V135G/V142G above, however, different RD models need to be used that better describe slow exchange processes.[108-109] The model used here to describe slow exchange RD data was developed by Lewis Kay's group and is labeled TSMFK01 from this point on. Parameters from this model can be compared to rate estimates from ZZ exchange and measured chemical shift differences. While chemical shift differences agree with measured values between exchanging pairs, RD rates are systematically higher than ZZ estimates. This result and additional evidence from the V145G variant prompted consideration of multi-state exchange models for unfolding of the LID domain.

The simplest multi-state exchange model consistent with the data is a 3-state parallel unfolding process. In this model a single folded, native conformation can unfold to the LU conformation by two pathways, one slow and the other intermediate-to-fast on the NMR timescale. This model is supported by V135G NMR data and unconfirmed evidence of slow exchange in HSQC spectra from the V142G and V135G/V142G variants. A number of studies have characterized multi-state exchange by NMR, however, more data than was collected in this study is required to get quantitative information about the states involved and rates of exchange.[88, 90, 108, 110-112] Quantitative analysis of multi-state exchange has been possible when the rates of the individual processes are on different NMR timescales. For instance, slow exchange processes can usually be quantified by ZZ exchange or CEST (conformational exchange saturation transfer), and subtracted from multi-state exchange RD data leaving the contribution of a faster process to be independently analyzed. [108] Multi-state exchange has also been quantified by collecting RD data as a function of temperature and including the enthalpy and entropy of the conformational changes in global fits of the data.[88] Another example characterizing multi-state exchange involved complex numerical models that require data from multiple field strengths, nuclei, and coherences.[90, 113] While multi-state

exchange is most consistent among all variants, a two-state analysis can give insight into how mutations promote the LU state. The complication of an additional exchange process may be minimal if one exchange process is dominant and contributes most to the RD signal recorded. The fold change in rates can be obtained by constructing a double mutant cycle, enumerating all known rates, and dividing the rate in question of the final by the initial state. This analysis could be flawed if the two mutation sites affect different exchange processes or if both processes contribute significantly to the observed exchange data. Altogether these data support a multi-state exchange, parallel unfolding model for LID local unfolding. Elucidation of this model would not be possible without the degenerate nature of the entropy-enhancing mutation scheme used here. In general, this mutation scheme should be able to obtain a deeper understanding about conformational dynamics involving high-energy intermediate states in proteins.

4.3 Materials & Methods

NMR ZZ Exchange

Theory

The modulation of the cross and self-peak intensity has been well described analytically for the 2-state process $F \Leftrightarrow U$ as follows:[106]

$$I_{nn}(t) = I_n(0) \cdot \frac{-(\lambda_2 - a_{11})e^{-\lambda_2 t} + (\lambda_1 - a_{11})e^{-\lambda_2 t}}{(\lambda_1 - \lambda_2)} \tag{11}$$

$$I_{uu}(t) = I_u(0) \cdot \frac{-(\lambda_2 - a_{22})e^{-\lambda_1 t} + (\lambda_1 - a_{22})e^{-\lambda_2 t}}{(\lambda_1 - \lambda_2)} \tag{12}$$

$$I_{nu}(t) = I_n(0) \cdot \frac{a_{21}e^{-\lambda_1 t} - a_{21}e^{-\lambda_2 t}}{(\lambda_1 - \lambda_2)} \tag{13}$$

$$I_{un}(t) = I_u(0) \cdot \frac{a_{12}e^{-\lambda_2 t} - a_{12}e^{-\lambda_2 t}}{(\lambda_1 - \lambda_2)} \tag{14}$$

where λ refers to the eigenvectors that describe the set of equations:

$$\lambda_{1,2} = \frac{1}{2}[(a_{11} + a_{22}) \pm [(a_{11} - a_{22})^2 + 4k_{nu}k_{un}]^{\frac{1}{2}}] \tag{15}$$

and a_{11}, a_{22}, and a_{12}:

$$a_{11} = R_{1,n} + k_{nu} \tag{16}$$

$$a_{12} = -k_{un} \tag{17}$$

$$a_{21} = -k_{nu} \tag{18}$$

$$a_{22} = R_{1,u} + k_{un} \tag{19}$$

Cross peak intensities, described by I_{nu} and I_{un} are predominantly dependent on the individual exchange rates at short mixing times resulting in a near linear buildup of cross peak intensity until spin-lattice relaxation, either $R_{1,n}$ or $R_{1,u}$ in this case, begins to take over. Exchange rates can be estimated by extracting the linear portion a polynomial fit to this buildup curve.

Experiment & Analysis

^{1}H-^{15}N ZZ exchange spectra were collected with mixing times of 0, 10, 30, 50, 70, 100, 150, 200, 250, 400, 700, and 1000 ms at 22 °C at 18.8 T with the 100 ms experiment run in triplicate and mixing times of 0, 10.6, 31.8, 53.1, 74.3, 106.1, 159.2, 222.8, 392.6, 700.4, and 1008.1 ms at 19° C at 14.1 T with the 106.1 ms experiment run in duplicate. All spectra were processed using nmrPipe, assignments transferred, and peak lists exported with the CCPNMR Analysis program. Peak list data were fit using Levenberg-Marquardt nonlinear-least-squares fitting implemented in the *lmfit* python package with equations 11-18 above. Rate constants were also estimated by fitting the first 5 mixing times for both 600

and 800 MHz data with a 3rd order polynomial and extracting the linear coefficient of the fit. Data and fits were plotted using ipython and matplotlib.

NMR Relaxation Dispersion

The TSMFK01 model that was used to analyze V135G RD data is described by the following equation:

$$R_{ex} = \frac{k_{AB} k_{BA} \sin(\Delta\delta \cdot \tau_{CPMG})}{\Delta\omega \cdot \tau_{CPMG}} \quad (20)$$

where k_{AB} and k_{BA} are the individual exchange rates, $\Delta\delta$ the chemical shift difference between the two states, and τ_{CPMG} ($1/\nu_{CPMG}$) the delay time between applied spin echo pulses during the NMR experiment. Both the CR72 and TSMFK01 models were fit using the *relax* analysis software and the TSMFK01 model chosen based on the Aikake information criterion. Analyses were repeated using in-house ipython scripts that imported the TSMFK01 model from *relax*. This was done because of limitations of the *relax* software. The *scipy* and *lmfit* python packages were used to perform the same grid search and Levenberg-Marquardt nonlinear-least-squares fitting as in *relax* and the analyses for V142G and V135G/V142G.

HSQC Experiments & 2-state Analysis

TROSY HSQC spectra were collected on the V135G/V142G variant at 2, 5, 7, 10, and 15 °C at 18.8 T to investigate evidence of slow exchange peaks in the vicinity of the folded conformation as observed in V135G and V142G. Data were processed and analyzed in the usual way described above.

The fold change in folding and unfolding rates were investigated assuming a two-state exchange model as follows:

$$\Omega_{F/U} = \frac{k_{F/U,Final}}{k_{F/U,Initial}} \quad (21)$$

where the first subscript identifies the rate, folding (F) or unfolding (U), and the second subscript identifies the protein. An asterisk distinguishes mutations made in a single variant from those in the WT. For instance:

$$\Omega^*_{F,135} = \frac{k_{F,DUB}}{k_{F,142}} \tag{22}$$

describes the change in the folding rate due to the valine to glycine mutation at position 135 in the context of the V142G variant. The rest of the mutation effects are as follows:

$$\Omega^*_{U,135} = \frac{k_{U,DUB}}{k_{U,142}} \tag{23}$$

$$\Omega^*_{F,142} = \frac{k_{F,DUB}}{k_{F,135}} \tag{24}$$

$$\Omega^*_{U,142} = \frac{k_{U,DUB}}{k_{U,135}} \tag{25}$$

Individual rates for the V142G and V135G/V142G variants were calculated from the rate constant and major state population fit parameters from the RD data.

CEST Theory

Chemical exchange saturation transfer, or CEST, is an alternative method for investigating conformational exchange in macromolecules by NMR. In these experiments, weak selective pulses are applied to varying frequencies in the dimension of interest, in the case of the HSQC-based experiment the weak fields are applied along the ^{15}N dimension. During the experiment a mixing time is placed between labelling the ^{15}N nucleus and the ^{1}H nucleus, similar to the ZZ-exchange experiment described in section, wherein the magnetization is placed along the z-axis or in line with the static magnetic field. During this mixing time the weak radio frequency pulse is applied perpendicular to the magnetization causing it to

rotate away from the z-axis. When magnetization is transferred back to the x-y plane and later detected any nuclei with spins in the range of the applied weak field have severely decreased peak intensity; however, nuclei with minor states at the frequency of the applied field will also have decreased signal intensity owing to chemical exchange with the minor state, thereby decreasing the magnitude of the magnetization. This weak field is varied throughout the ^{15}N spectral width and the intensity of the visible peak is observed as a function of the frequency of the applied field.

4.4 Results & Discussion

Measured Conformational Exchange in V135G

V135G is in clear slow exchange between the native and LU conformations and ZZ exchange can be used to quantify the rate of exchange between these states. The ^{1}H-^{15}N ZZ exchange experiment is essentially an HSQC with an added mixing time that allows conformational exchange to occur and results in the appearance of cross peak intensity. As the mixing time is increased the two cross peak intensities increase linearly until longitudinal relaxation begins to diminish peak intensity. ZZ exchange curves were collected at both 22 and 19 °C on 18.8 and 14.1 T spectrometers, respectively. Data for many residues at 14.1 T were not available because of peak broadening and overlap, therefore 18.8 T data were used for analysis. Two-state analytical exchange models were not able to fit the data as seen in Figure 4.1. Therefore, rates were estimated from the linear region of cross peak intensity by fitting a 2^{nd} order polynomial to the first 5 mixing times of normalized cross peak intensities, shown in Figure 4.2. Estimated rates from the linear term of the polynomial cluster around 2 and 12 s^{-1} for folding and unfolding rate, respectively, as seen in Figure 4.8. These rates are an order of magnitude lower than the exchange rates measured in the other variants. Due to peak overlap in ZZ exchange spectra and the inability to fit the ZZ exchange data, RD data were collected on V135G at 19 °C and analyzed separately for native and LU sets of peaks. The TSMFK01 model, developed for slow exchange data, was used to fit the data. The native

and LU data, shown in Figure 4.3 and Figure 4.5, respectively, were fit well by the model which recapitulated the oscillatory behavior that is a hallmark of slow exchange data. Fits to the folded set of data in Figure 4.4 produced large chemical shift differences and unfolding rates that cluster about a median value of 19 s^{-1} for LID residues. Fits to the LU data produced folding rates clustering near 5 s^{-1} within the LID and large chemical shift differences consistent with observed values. These folding and unfolding rates from RD fits are on the same order of magnitude as the ZZ exchange estimates but are systematically larger as seen in Figure 4.8. The magnitude of chemical shift differences from fits to both sets of RD data correlate well with measured values between the slow exchange pairs as seen in panel A of Figure 4.7. The RD parameters are consistent with the observed local unfolding conformational change in the LID on a slow-exchange timescale.

Although the RD data for V135G are consistent with a slow, two-state process, systematic differences in fits to the data indicate an additional exchange process. Fits to the native dataset are not as good as those to the LU data. Although in Figure 4.7 A reduced χ^2 values are constant for fits to the LU dataset, χ^2 values from the native dataset are higher and scale with increasing chemical shift differences. Furthermore, $\Delta\delta$ parameter values are systematically larger than measured values, and scale with the magnitude of the chemical shift difference, as seen in Figure 4.7 B. These results are consistent with an additional exchange process primarily affecting the native state but accessing the same LU state on a faster timescale of exchange. Such a process would shift the observed native peaks closer to their corresponding LU peaks resulting in larger $\Delta\delta$ fit parameters versus observed values that scale with the magnitude of the chemical shift differences. Furthermore, additional exchange would lead to greater R_{ex} values and cause rates of exchange from RD fits to be larger, by equation (20), than those from ZZ exchange estimates as observed in Figure 4.8. To further test the two-state assumption for the V135G data a van't Hoff analysis was performed with NMR data. An added benefit of the presence of peak intensity for both the native and LU populations of the LID domain in V135G is the ability to observe populations modulate with temperature. Previous thermodynamic studies predicted a large change in the LU population over a narrow temperature range. As expected peak intensity shifts from native to LU peaks with increasing temperature

from 10 to 25 °C as seen in Figure 4.9, consistent with the expected behavior for local unfolding. However, peak intensity decreases for both native and LU peaks below 19 °C. The behavior of the peak intensity directly contradicts the expected trend. The loss in native peak intensity could be the result of these peaks moving into intermediate exchange or the native equilibrium population shifting to a third, unknown state. Regardless, peak intensities collected at 19, 22, and 25 °C were used to perform a van't Hoff analysis on a per residue basis. The enthalpies and melting temperatures are quite variable and lowest within the LID domain as displayed in Figure 4.11. The enthalpies are quite large, which is consistent with local unfolding, however, both enthalpies and melting temperatures were lower than those determined by differential scanning calorimetry. Based on this analysis LID unfolding in V135G is not two-state.

3-state model consideration

With the data so far, speculation can be made about the multi-state exchange process. It appears that the additional exchange process primarily affects the folded set of peaks. The simplest model therefore is a 3-state exchange scheme in which molecules can locally unfold by either a fast or slow pathway and that both exchange processes share common structural endpoints:

$$B \rightleftharpoons A \rightleftharpoons C \tag{26}$$

where state A is the native LID conformation and states B and C are locally unfolded conformations with similar structural characteristics but different exchange kinetics. States B and C may be able to exchange, however, this process would not be directly observable by NMR. Furthermore, it appears that the different mutation sites affect different exchange processes with V142G promoting the fast exchange process and V135G the slow exchange process. The convolution of fast and slow exchange may explain the failure of the two-state model to fit to the V135G ZZ exchange data. The effects of an additional intermediate exchange process on a slow-exchange process has not been explored by the author or found in the literature. Multi-state exchange could also contribute to the discrepancy in the V135G LU equilibrium

population observed by NMR and determined from calorimetry as folded peaks in intermediate exchange could have significantly decreased peak intensity.

V142G and V135G/V142G both show signs of slow exchange in addition to RD data that is in fast-to-intermediate exchange. Figure 4.13 highlights peaks in the V142G spectrum collected at 19 °C that overlay with the LU peaks identified in both V135G and V135G/V142G. Conversely, peaks at folded LID positions, circled in Figure 4.14, for many of the same residues appear in the V135G/V142G variant at temperatures below 7 °C. RD data from these same residues are in intermediate-to-fast exchange in both variants as compared to calculated coalescence rates shown in Figure 4.15. The exchange rate in the V135G/V142G variant could have moved into slow exchange with the decrease in temperature, however, this is inconsistent with the persistence of LU peak intensity as temperatures decrease as these peaks would disappear if exchange moved from fast to slow. The intensity of the purported slow exchange pairs is too low to be verified by ZZ exchange methods, however, the fact that multiple peaks show this behavior suggests that these are most likely slow exchange pairs. Furthermore, multi-state exchange could lead to the observed variability in exchange rates, major state populations, and systematic differences in chemical shift differences determined from RD experiments in V142G and V135G/V142G.

Inferences can be made about the nature of the different exchange processes. Slower folding occurs in proteins with large absolute contact order.[114] The LID region containing residues 110-180 has a contact order of 6.4 and 5.9 for the open and closed crystal structures, respectively. This places the folding rate of the LID domain itself within 300-400 s^{-1}, which is consistent with exchange rates determined from V142G and V135G/V142G. The slow exchange process could be influenced by proline isomerization; however, the rate of slow exchange is much faster than measured rates of proline isomerization. There are three proline residues in the LID domain at positions 128, 139, and 140. The two adjacent prolines are also conserved amongst AKs that contain a LID and appear to stabilize residues F137 and N138 in the LID that interact with the bisubstrate inhibitor AP5A. It is unclear what role, if any, these prolines may play in the local unfolding process. The presence of one or more *cis* prolines in the LID domain could prevent efficient folding and reduce the observed exchange rate. Removal of one or more of the three prolines in

the LID domain could test this hypothesis. Ideally such mutations would eradicate the slow exchange process allowing the fast process to be characterized.

Multi-state exchange in AK can be characterized by NMR. As stated above many studies have characterized multi-state exchange. However, with the current amount of data such analyses are not possible. A major obstacle is the inability to fit the ZZ exchange data from V135G and the lack of slow exchange rates in the other variants. Slow exchange measurements could be made using CEST NMR experiments described in the materials and methods. In short, CEST scrambles magnetization at defined ^{15}N frequencies decreasing intensity of peaks at that frequency and peaks in different regions of the spectrum with exchange pairs at that frequency. In this way, peaks with weak intensity or spectra that are too crowded for ZZ exchange experiments can be characterized. This is a perfect experiment for the lowly-populated slow exchange pairs for both V142G and V135G/V142G; however, the influence of the intermediate exchange process in CEST experiments is unknown and could complicate the data analysis. Lastly, RD data could be collected at multiple temperatures or on multiple nuclei and coherences to better resolve a three-state analysis which require more complicated analytical and numerical models.[113] These experiments pose difficulties since the analysis of the temperature data is specialized and not implemented in many RD analysis software packages including *relax* and pulse sequences for collecting RD data on multiple nuclei and coherences are not implemented at the JHU NMR facility.

2-state analysis

Since quantitative analysis of multi-state exchange is not currently feasible data from all AK variants was analyzed with a two-state exchange model. The analysis is illustrated by the thermodynamic cycle in Figure 4.12. The effect of each mutation was considered separately for the folding and unfolding rates and depicted by Ω terms as defined in equation (21) above. Effects were further separated by the context in which the mutation was made, in the WT protein or a single variant. Only rate changes between a single and double mutation could be considered since data were not able to be collected for the LU

process in the WT protein. From this analysis, the two mutation sites have different effects on the folding and unfolding rates as seen in Table 4-3. Stabilization of the LU conformation can occur because of reductions in the folding rate, increases in the unfolding rate, or a combination of both. The V135G mutation stabilizes the LU state increasing the unfolding rate by 12-fold and decreasing the folding rate by 30-fold. On the other hand, the V142G mutation increases both rates and stabilizes the LU state through the unfolding rate only. Cooperativity between the mutation sites could be a reason for the differential effects that mutations have on the rates of exchange. However, without the WT exchange rates the effects of single mutations on their own cannot be determined and therefore cooperative between multiple mutations cannot be assessed. Another reason for the different actions of the mutations could be due to effects on the local sequence, such as nearby prolines, and partial packing of side chains of mutated valines.

4.5 Conclusions

The V135G variant is in slow exchange between the native and LU conformations. Despite being in slow exchange analytical models cannot fit the ZZ exchange data. RD data verify that slow exchange is occurring on the order of 20 s^{-1}. The disagreement between van't Hoff enthalpies from NMR data and previous calorimetric values and the loss in folded peak intensity at low temperatures provides strong evidence for multi-state exchange. A three-state exchange model with a single native state and two, structurally identical unfolded states with different rates of exchange is most consistent with the data. In accordance with this model both V142G and V135G/V142G show evidence of slow exchange pairs in addition to the measured fast-exchange process. The intermediate-to-fast exchange process is on the order of 300-400 s^{-1}, in accordance with estimates based on contact order, and the slow process occurs at 20 s^{-1}. Current data are insufficient to definitively characterize this process. Regardless, this study shows the complexity of a process as simple as the unfolding of a single domain of 60 residues and that entropy-enhancing mutations can illuminate processes that are otherwise unobservable. These mutations may prove invaluable in the study of high-energy conformations in other proteins.

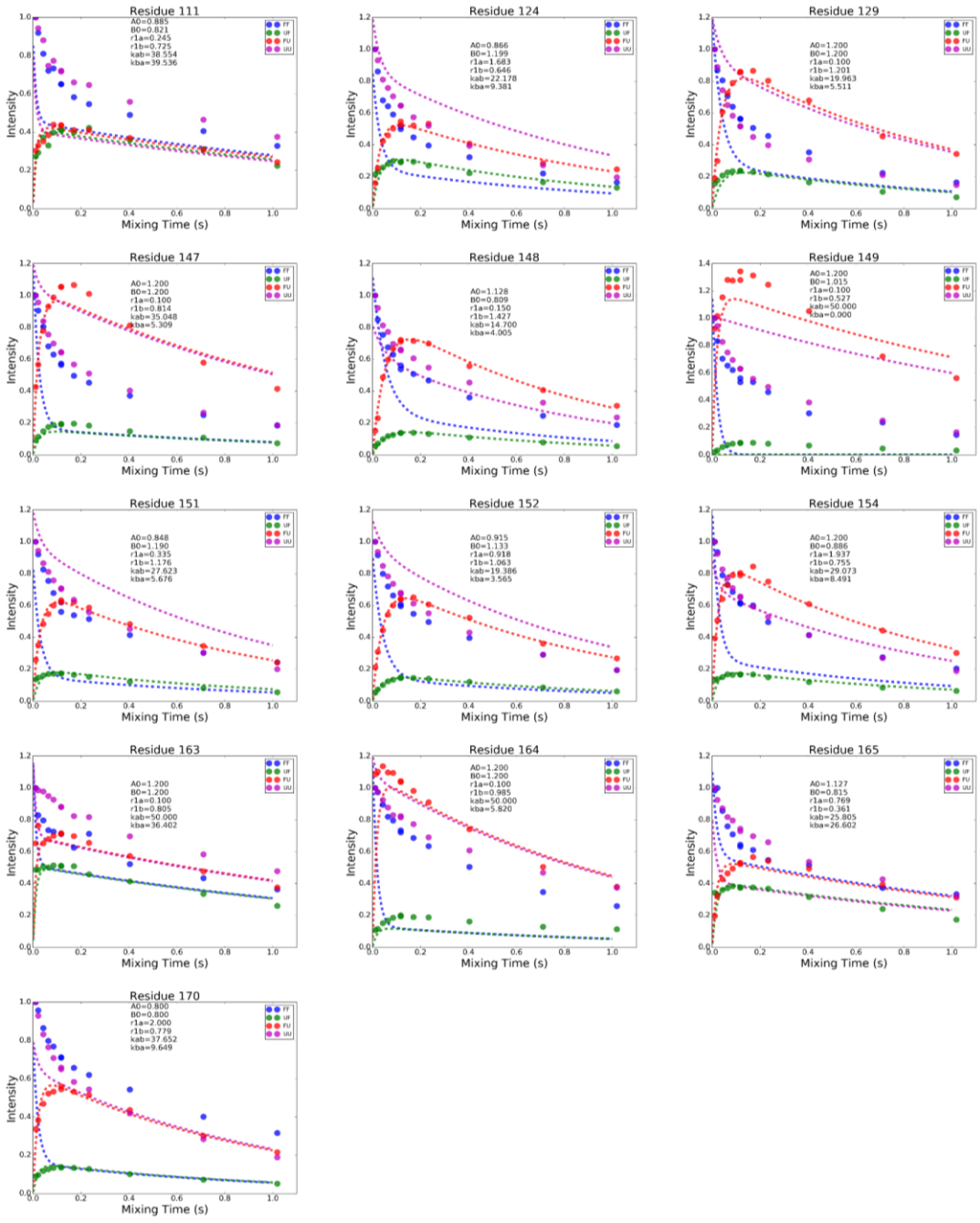

Figure 4.1 V135G ZZ Exchange Rate Estimation

ZZ exchange data collected at 22 °C, a pH of 7.0, and 18.8 T were not able to be fit with a 2-state exchange model. Under these conditions the observed equilibrium constant (N/LU) is 0.4. The fits above are to cross peak intensities only. Data are as points and fits as dashed lines.

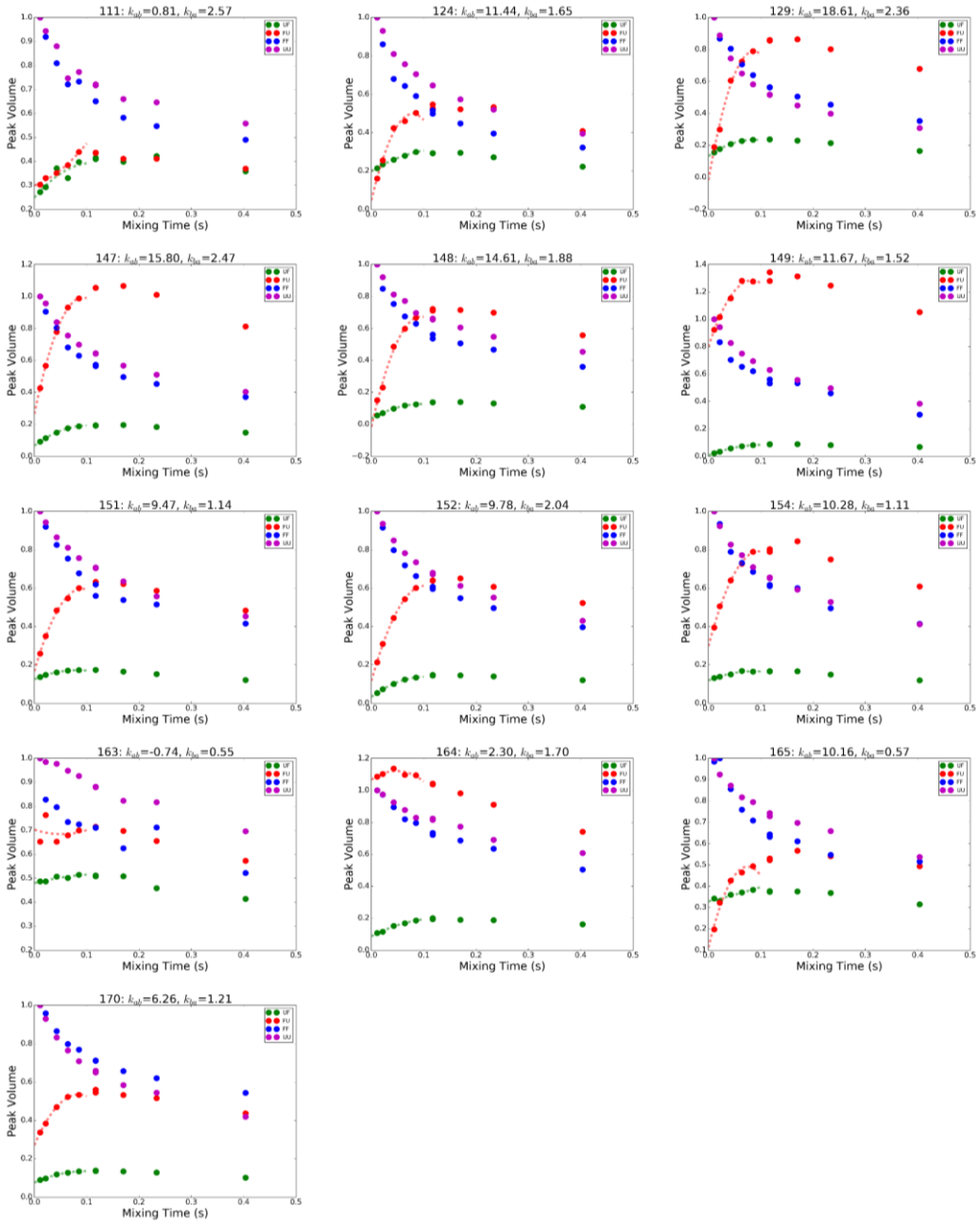

Figure 4.2 V135G ZZ Polynomial Fits

Estimates of exchange rates can be made using the linear portion of cross peak buildup curves. Shown above are 3rd order polynomial fits to the first 5 data points of cross peak buildup curves for ZZ exchange data collected at 22 °C, a pH of 7.0, and 18.8 T. The linear portion of these fits were used as estimates of exchange rates shown in Figure 4.8.

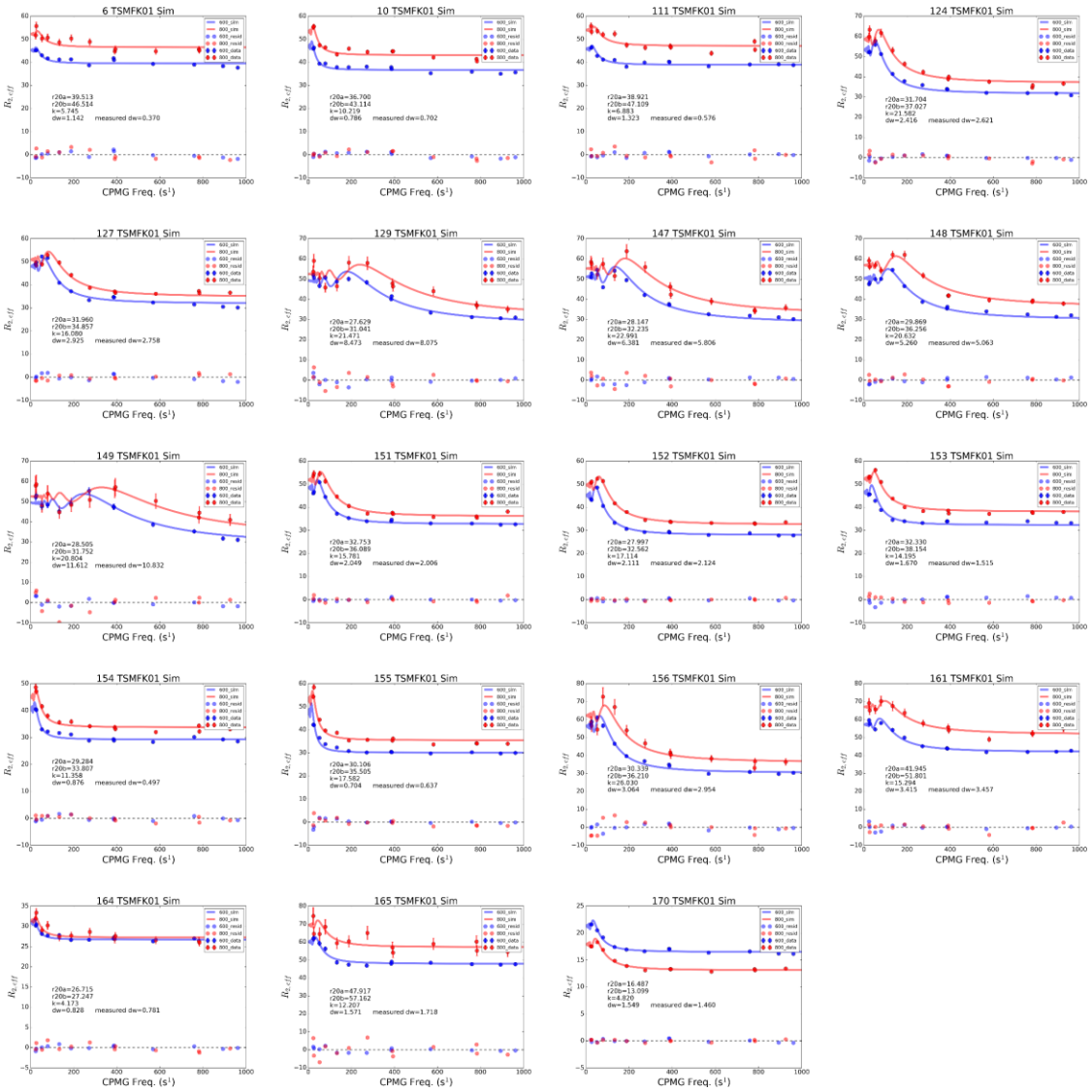

Figure 4.3 V135G Folded RD Curves

CPMG curves from the folded set of slow exchange peaks from the V135G variant show definite signs of exchange (points) and reasonable fits to the TSMFK01 slow exchange model (lines). Data were collected at 22 °C and 14.1 (blue) and 18.8 (red) T.

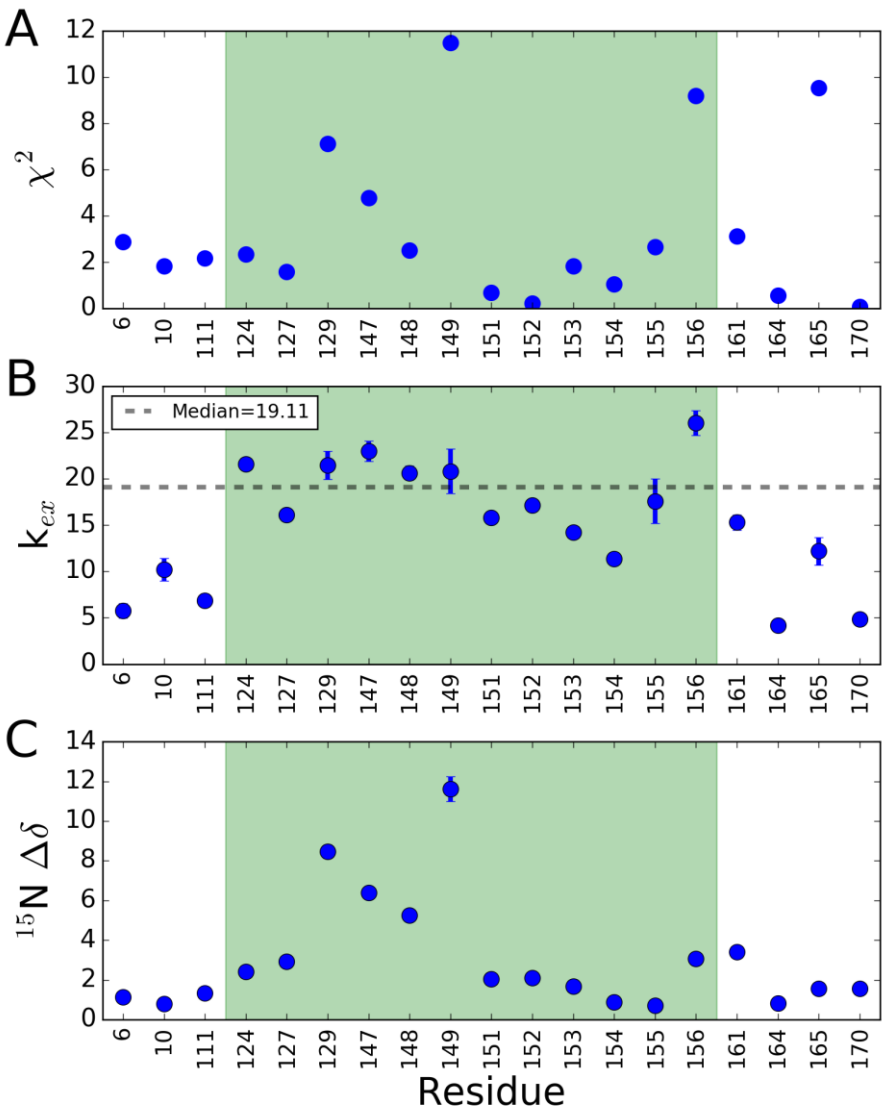

Figure 4.4 V135G FF TSMFK01 Fit Parameters

A. The reduced χ^2 values from fits of the TSMFK01 analytical model to CPMG data from the folded set of slow exchange peaks (shown in Figure 4.3) in the V135G variant show generally good fits for residues that have an estimated R_{ex} of at least 1 s^{-1}. B. Unfolding rates cluster about the median of 19.11 s^{-1} for residues in the LID (green). C. ^{15}N chemical shift differences are largest within the LID domain. Errors are derived from Monte Carlo simulations in all cases and the green region highlights LID residues.

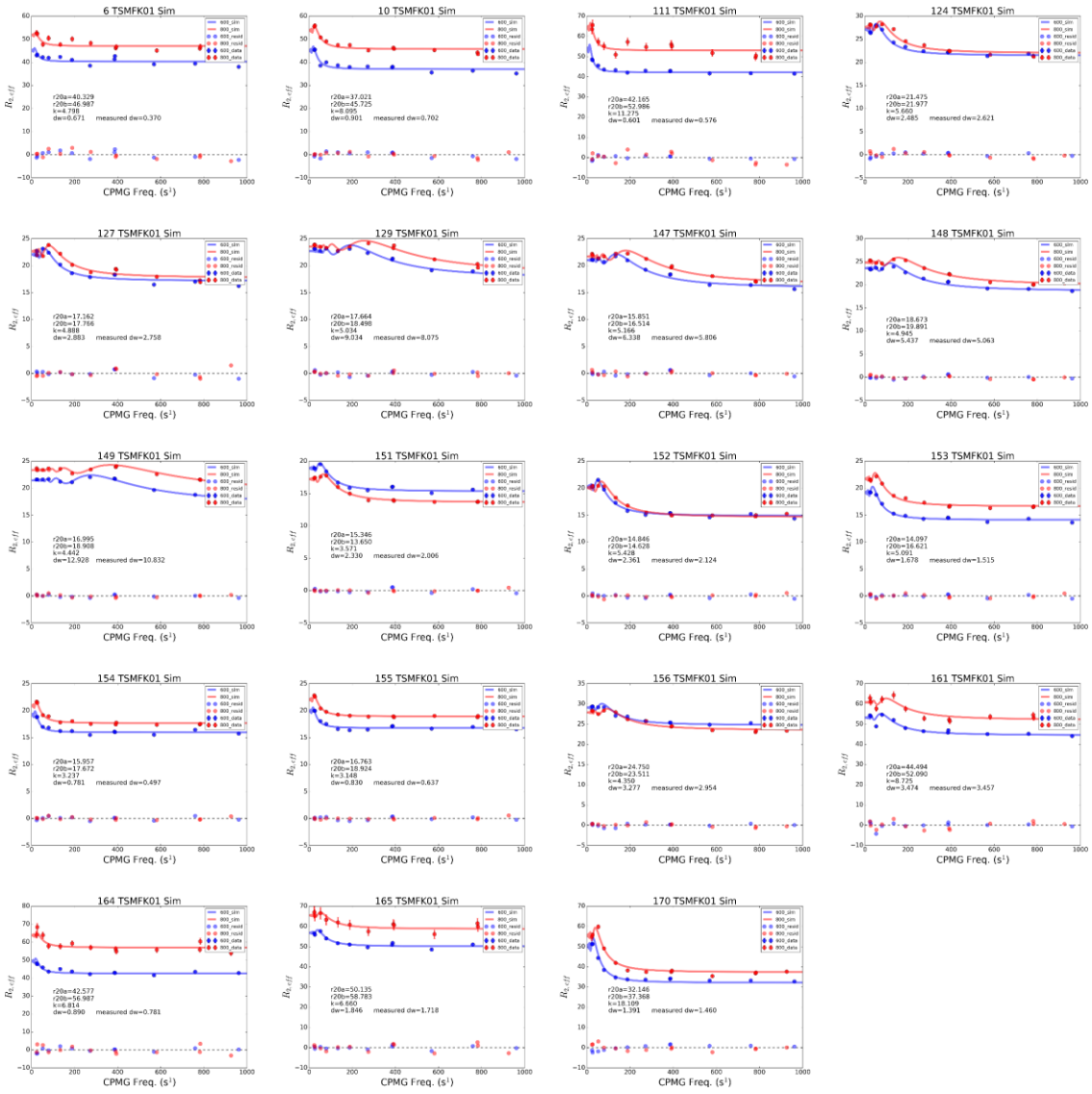

Figure 4.5 V135G Unfolded RD Curves

CPMG curves from the locally unfolded set of slow exchange peaks in the V135G variant show definite signs of exchange (points) and the TSMFK01 slow exchange model (lines) fits the data well. Data were collected at 22 °C and 14.1 (blue) and 18.8 (red) T.

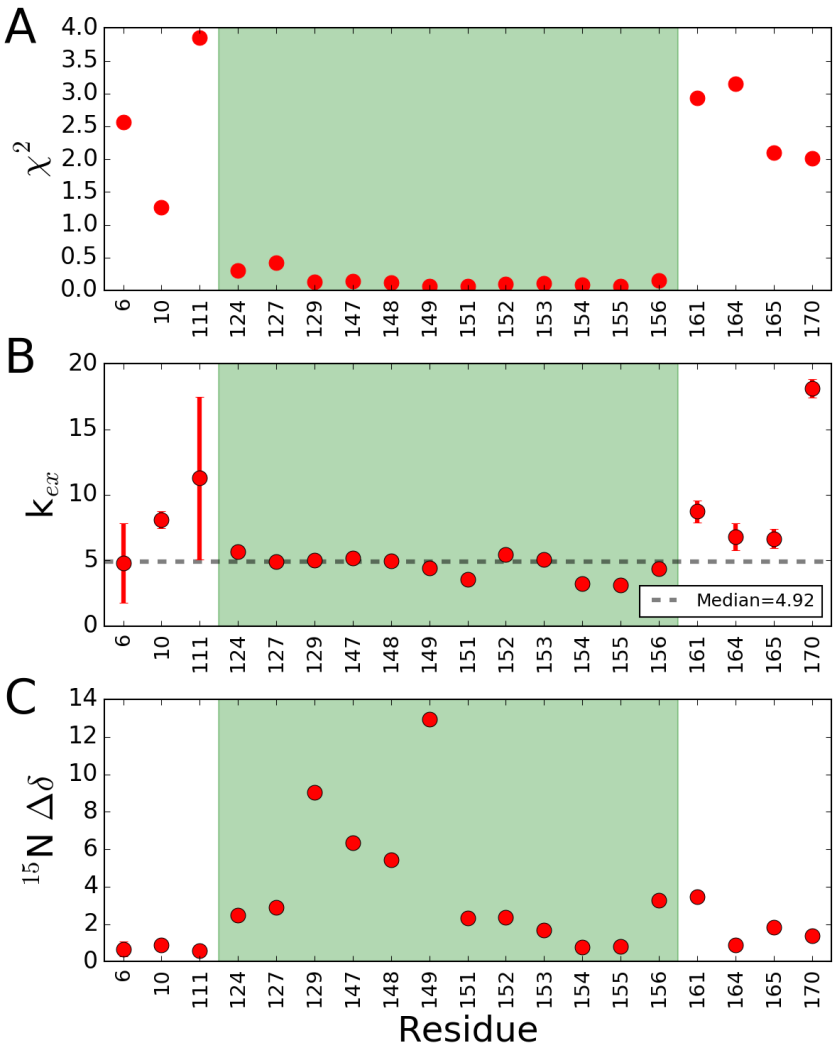

Figure 4.6 V135G UU TSMFK01 Fit Parameters

A. The reduced χ^2 values from fits of the TSMFK01 analytical model to CPMG data from the unfolded set of slow exchange peaks (shown in Figure 4.5) in the V135G variant show excellent fits for residues that have an estimated R_{ex} of at least 1 s^{-1}. B. Within the LID domain the folding rates cluster about the median of 4.92 s^{-1}. C. ^{15}N chemical shift differences are largest within the LID domain. Errors are derived from Monte Carlo simulations in all cases and the green region highlights LID residues

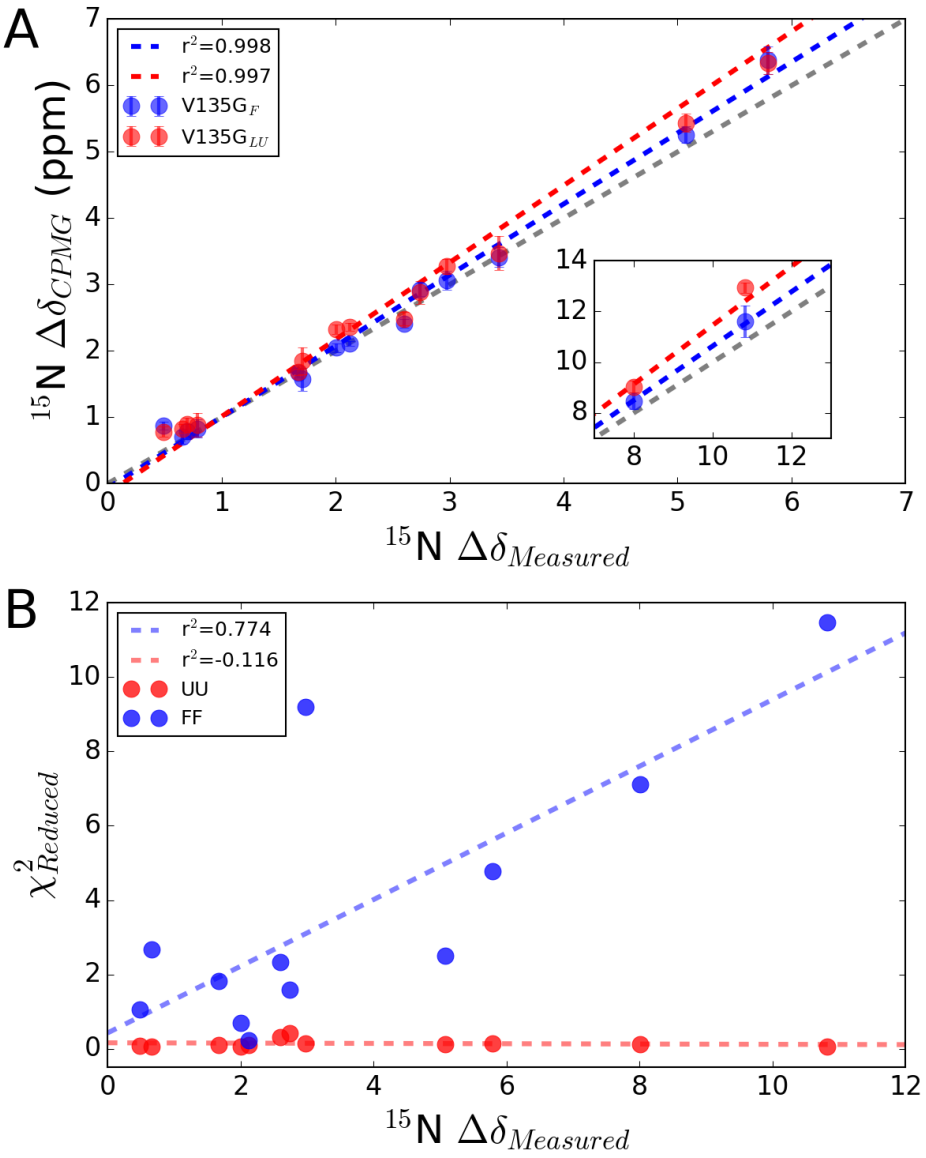

Figure 4.7 V135G $\Delta\partial_{fit}$ and χ^2 Compare

A. Measured and fitted chemical shift differences correlate well for both the folded and LU sets of data but both deviate from the diagonal at higher measured values. B. The LU fits are equally good across a large range of chemical shift differences whereas the folded fits are increasingly worse at larger shift differences. These data are consistent with a contribution of an additional exchange process that primarily affects the folded set of peaks since the R_{ex} contribution scales with the chemical shift.

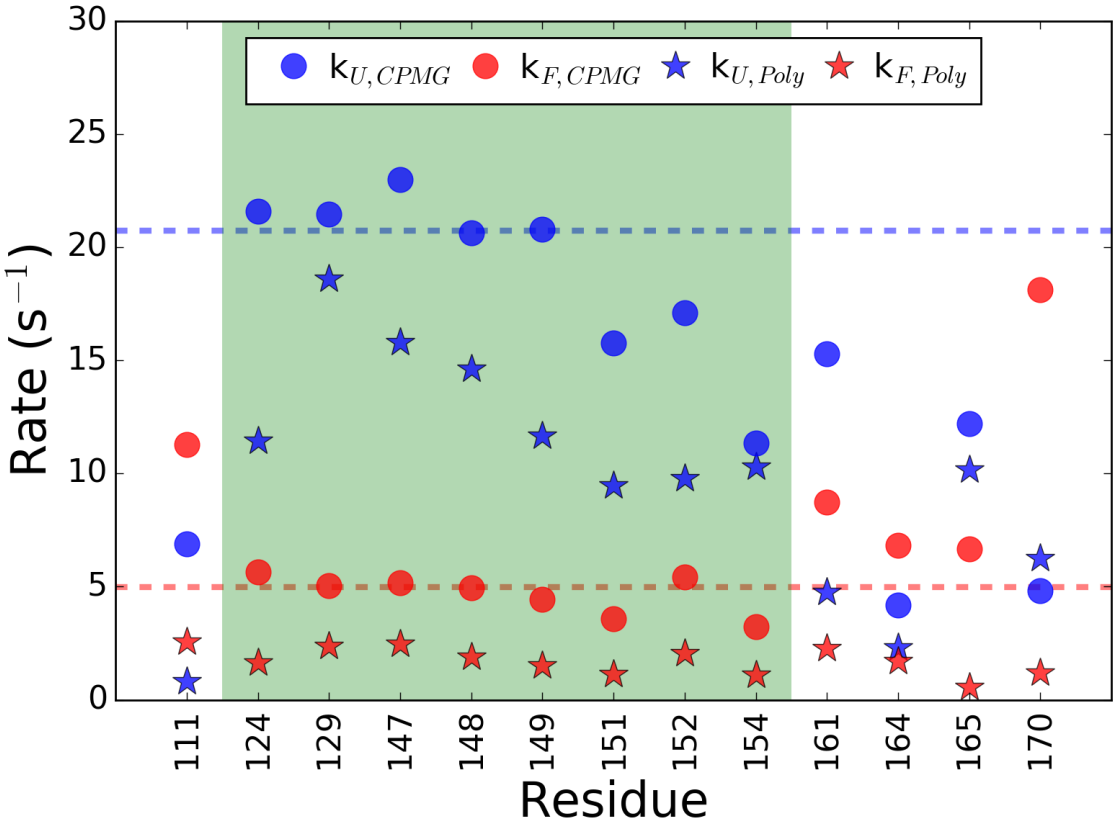

Figure 4.8 V135G ZZ/RD Rate Comparison

V135G folding and unfolding rates, in red and blue, respectively, of exchange determined from polynomial fits to the crosspeak buildup curves of the ZZ exchange data (Figure 4.2) shown as stars and those determined from fits to RD data using the TSMFK01 model in circles showing similar trends in the rates determined by both methods. Estimated ZZ exchange rates are systematically lower than those determined from CPMG data for both the folding and unfolding rates.

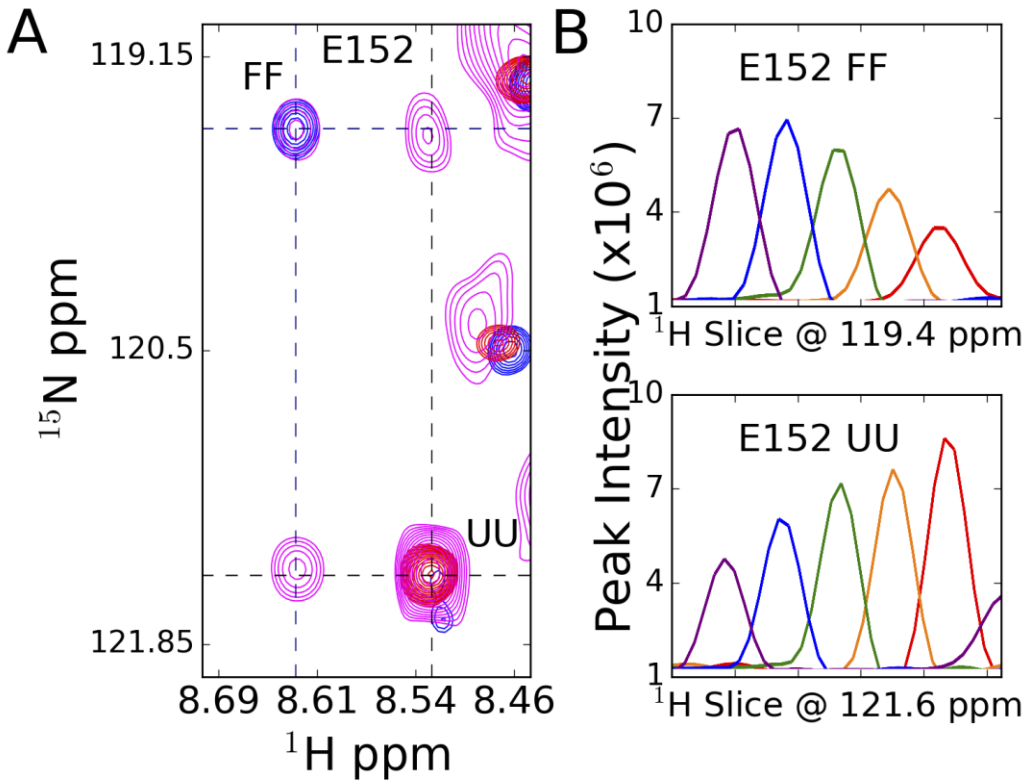

Figure 4.9 Evidence of Slow Exchange in V135G

A. ^{1}H-^{15}N ZZ exchange spectrum highlighting E152 with V142G in blue, V135G/V142G in red, and V135G in magenta showing peak intensity at folded, LU, and cross peak positions. B. Modulation of peak intensity with temperature for the folded (top) and unfolded (bottom) peaks of E152 in the V135G variant. Data were collected at 10, 15, 19, 22, and 25 °C at 18.8 T and a pH of 7.0 and shown in purple, blue, green, orange, and red, respectively.

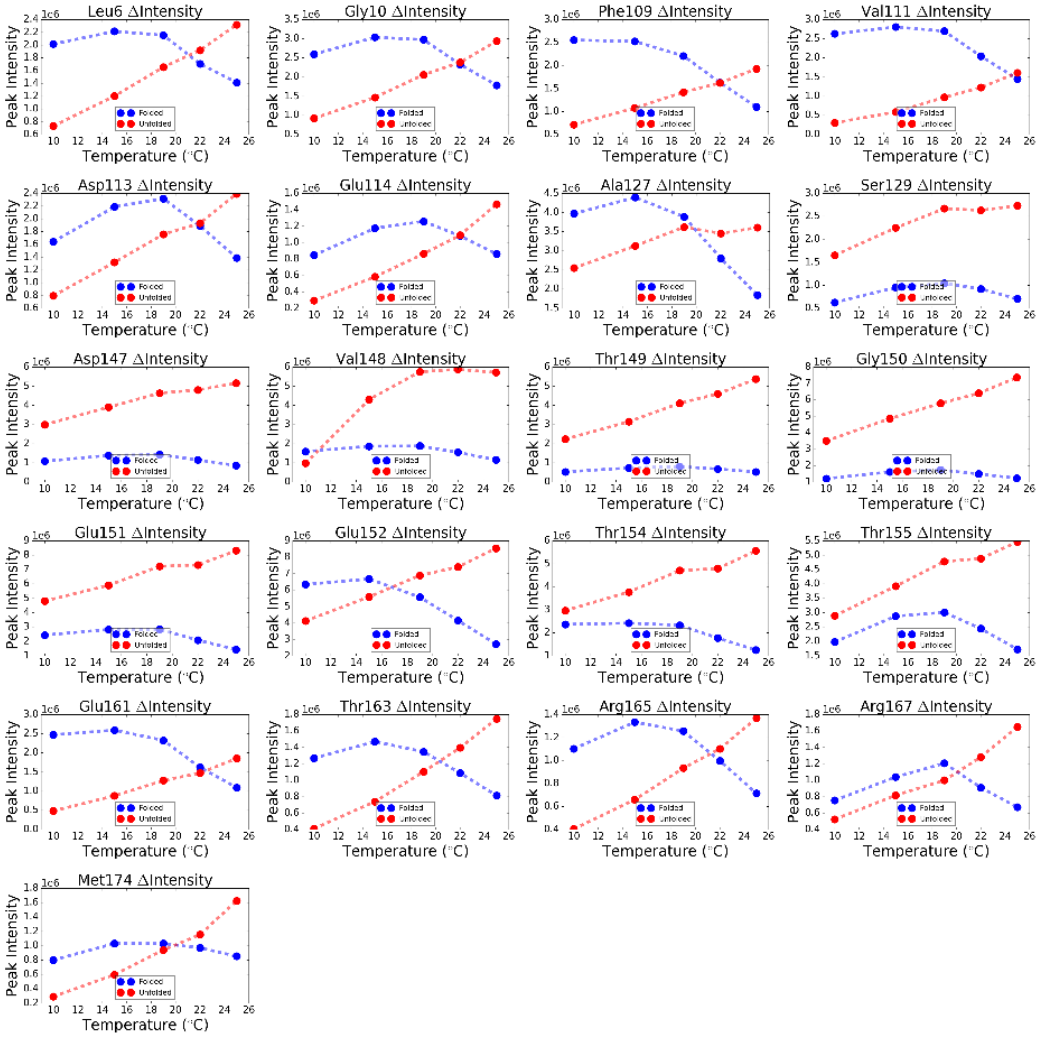

Figure 4.10 V135G Nonlinear Temperature Profiles

Intensity versus temperature profiles for slow exchange residues in V135G that have well dispersed peaks in both conformations. Folded peak intensity is diminished non-linearly at low temperatures while the unfolded peak intensity remains linear over all temperatures.

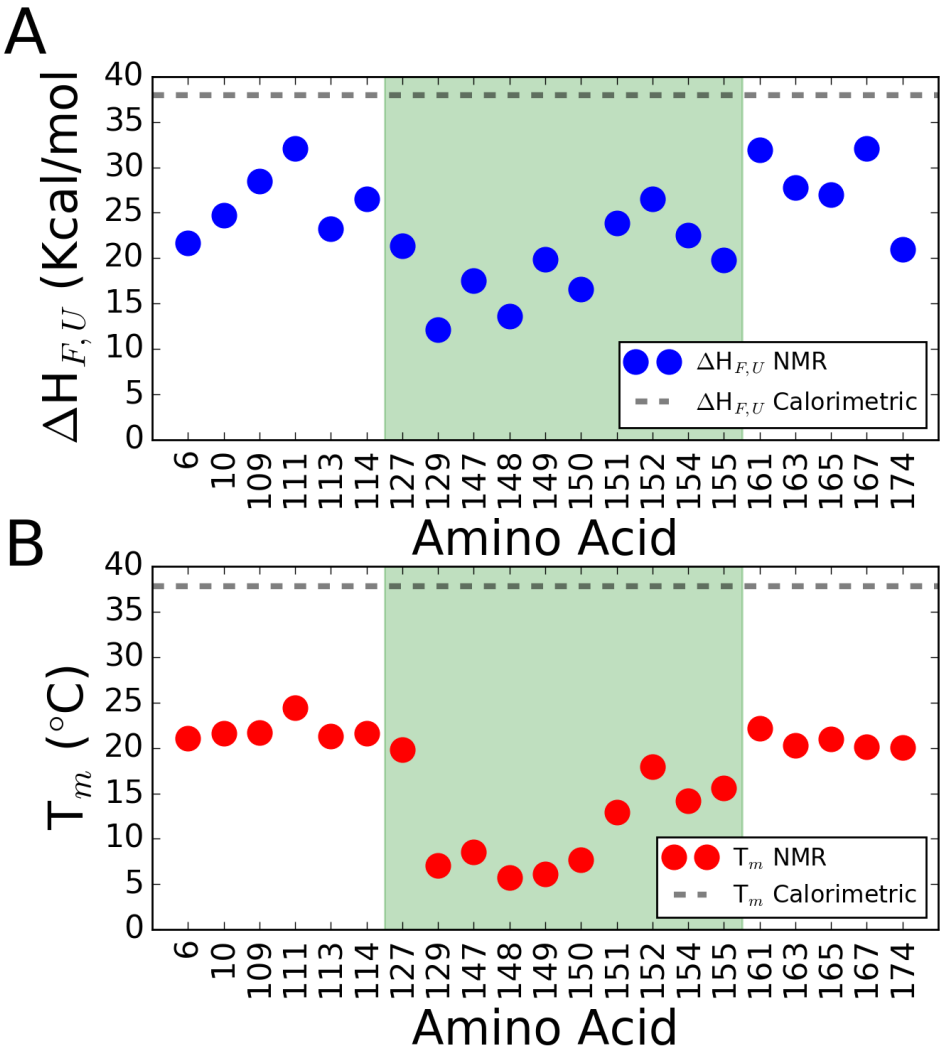

Figure 4.11 V135G LID van't Hoff Enthalpy

A. van't Hoff enthalpies calculated from slow exchange peaks in TROSY HSQC spectra collected at 19, 22, and 25° C. The green box indicates LID domain residues and grey dashed line represents the calorimetric value obtained previously by DSC. B. Corresponding melting temperatures from the van't Hoff analysis. Green box is the same as A and dashed line is the T_M from previous calorimetric measurements.

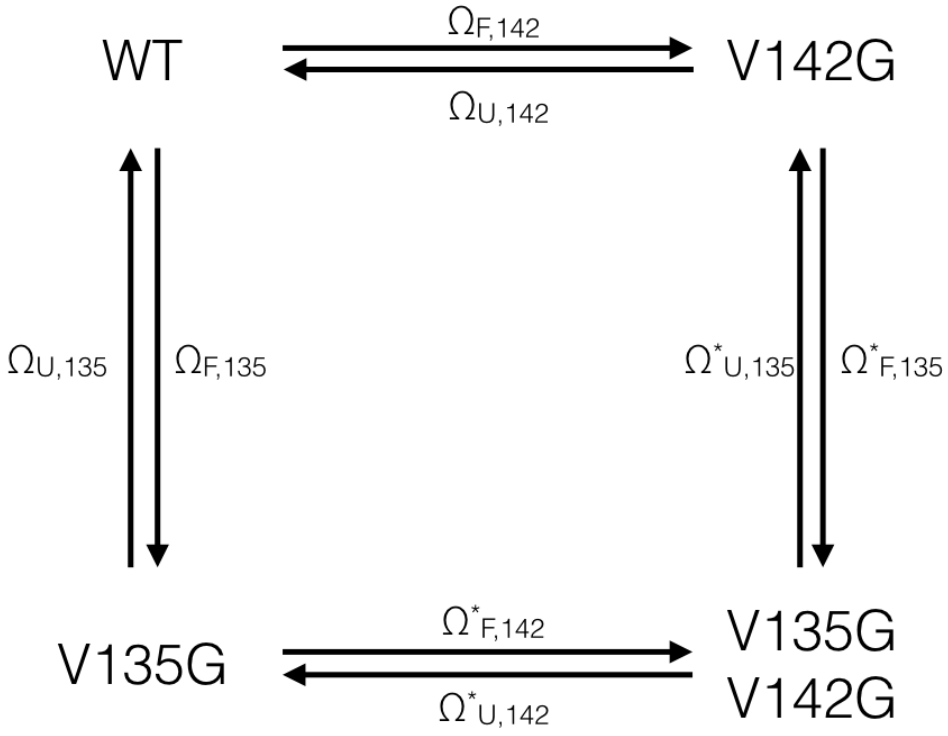

Figure 4.12 AK Variant Cycle

Thermodynamic cycle of AK variants in this study with the fold change in folding and unfolding rates depicted individually by the Ω terms. Astersisks distinguish fold changes upon a second mutation.

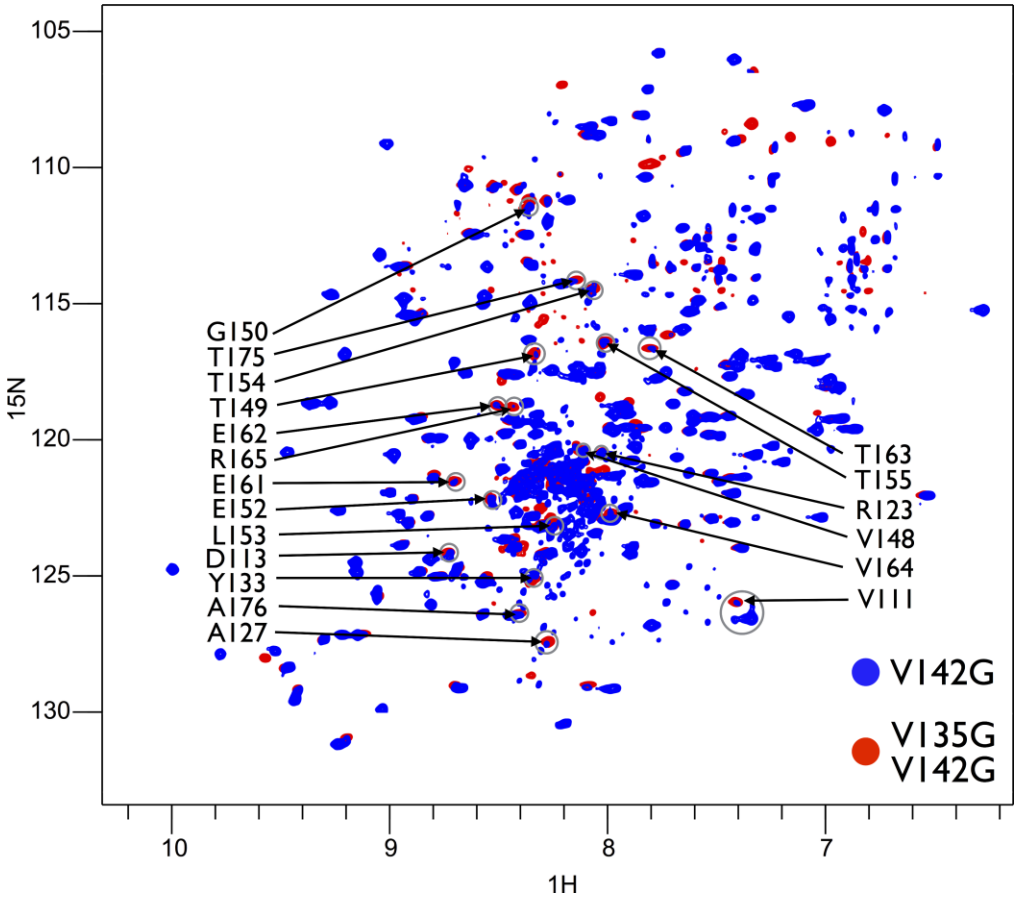

Figure 4.13 V142G Slow Exchange Overlay

Overlay of TROSY HSQC spectra from V142G (blue) and V135G/V142G (red) collected at 19°C. Circled and labeled peaks highlight assigned LID residues in V135G/V142G where unassigned peak intensity also appears in the V142G spectrum.

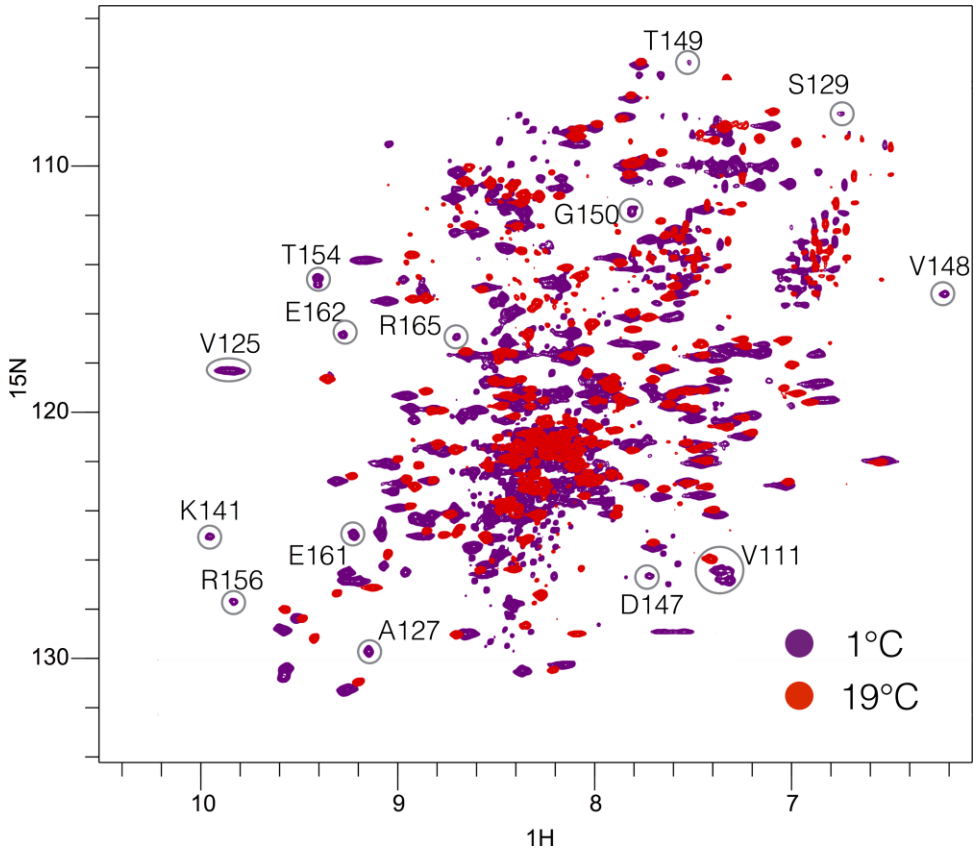

Figure 4.14 V135GV142G Slow Exchange Overlay

Overlay of TROSY HSQC spectra from V135G/V142G at 19°C (red) and at 1°C (purple). Circled and labeled peaks observable at 1°C overlay with native peak positions in the V142G and V135G variants and highlight evidence of slow exchange in V135G/V142G.

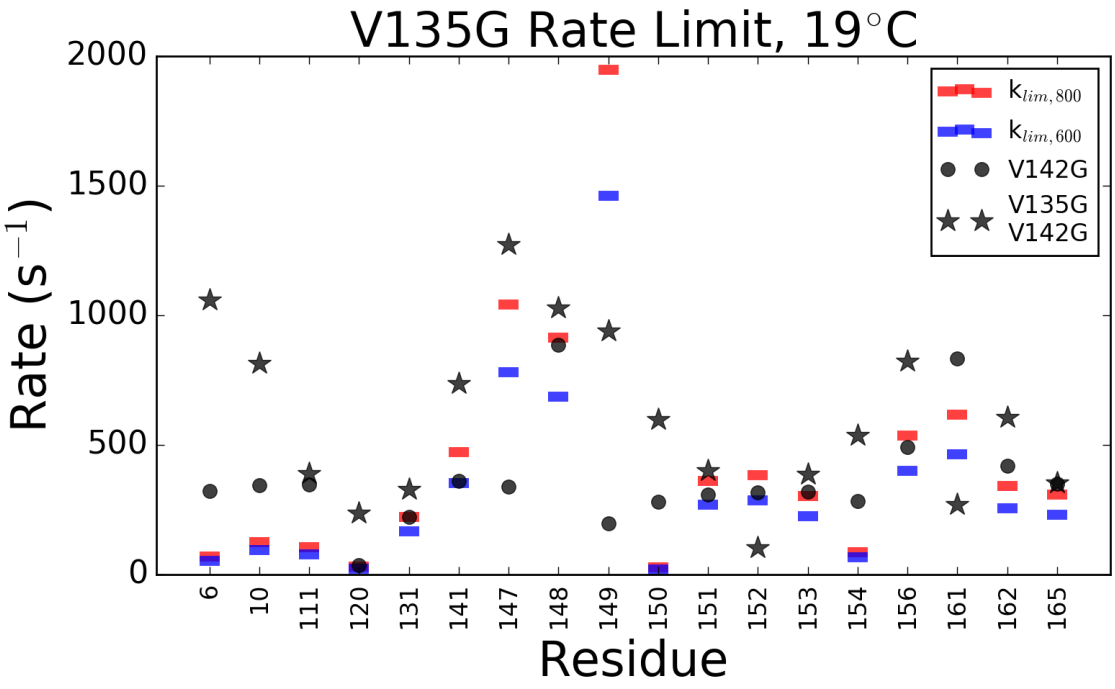

Figure 4.15 Slow Exchange Rate Limits

Convalescence rates calculated from measured $\Delta\partial$ values in the V135G variant at both 600 and 800 MHz field strengths as blue and red lines, respectively. The rate constants from RD data analysis for V142G and V135G/V142G are plotted as points and stars, respectively, showing that most residues are in intermediate to fast exchange.

Table 4-1 V135G UU TSMFK01 CPMG Fit Parameters

Parameters presented in Figure 4.6 from fits to CPMG data from the unfolded set of slow exchange peaks in the V135G variant shown in Figure 4.5.

	χ^2	$k_{fold}(s^{-1})$	error	^{15}N $\Delta\delta$ (ppm)	error
6	2.56	4.8	3.05	0.67	0.4
10	1.26	8.09	0.64	0.9	0.09
111	3.86	11.28	6.21	0.6	0.27
124	0.31	5.66	0.25	2.49	0.1
127	0.42	4.89	0.3	2.88	0.17
129	0.13	5.03	0.23	9.03	0.22
147	0.14	5.17	0.2	6.34	0.16
148	0.12	4.94	0.18	5.44	0.14
149	0.06	4.44	0.21	12.93	0.21
151	0.07	3.57	0.12	2.33	0.07
152	0.09	5.43	0.14	2.36	0.06
153	0.1	5.09	0.16	1.68	0.05
154	0.08	3.24	0.28	0.78	0.08
155	0.07	3.15	0.19	0.83	0.06
156	0.15	4.35	0.18	3.28	0.11
161	2.93	8.73	0.84	3.47	0.25
164	3.15	6.81	1.04	0.89	0.18
165	2.1	6.66	0.74	1.85	0.2
170	2.01	18.11	0.71	1.39	0.05

Table 4-2 V135G FF TSMFK01 CPMG Fit Parameters

Parameters presented in Figure 4.4 from fits to CPMG data from the folded set of slow exchange peaks in the V135G variant shown in Figure 4.3.

	χ^2	$k_{unfold}(s^{-1})$	error	^{15}N $\Delta\delta$ (ppm)	error
6	2.87	5.75	0.79	1.14	0.16
10	1.84	10.22	1.22	0.79	0.11
111	2.18	6.88	0.71	1.32	0.12
124	2.33	21.58	0.67	2.42	0.07
127	1.59	16.08	0.55	2.93	0.1
129	7.1	21.47	1.51	8.47	0.33
147	4.77	22.99	1.1	6.38	0.21
148	2.51	20.63	0.76	5.26	0.13
149	11.47	20.8	2.41	11.61	0.63
151	0.69	15.78	0.4	2.05	0.06
152	0.22	17.11	0.22	2.11	0.03
153	1.83	14.2	0.66	1.67	0.07
154	1.06	11.36	0.62	0.88	0.06
155	2.66	17.58	2.42	0.7	0.1
156	9.2	26.03	1.34	3.06	0.14
161	3.12	15.29	0.82	3.41	0.15
164	0.58	4.17	0.55	0.83	0.14
165	9.53	12.21	1.49	1.57	0.17
170	0.07	4.82	0.13	1.55	0.04

Table 4-3 2-state Mutational Effects

Fold change in folding or unfolding rates upon mutation as defined in equation (21).

	$\Omega^*_{135,F}$	$\Omega^*_{135,U}$	$\Omega^*_{142,F}$	$\Omega^*_{142,U}$
125	0.02	12.0	3.6	20.3
147	0.04	32.5	2.3	54.9
148	0.01	21.0	2.1	49.4
149	0.06	43.8	2.5	44.6
151	0.03	12.8	2.4	24.8
152	0.04	2.2	1.8	5.5
153	0.04	9.3	1.9	26.6
154	0.28	3.5	12.2	43.9
156	0.03	18.2	3.0	31.1
Median	0.04	12.8	2.4	31.1

Table 4-4 2-state Exchange Rates

Rates used in the calculation of the Ω values in Table 4-3.

	$k_{F,135}$	$k_{U,135}$	$k_{F,142}$	$k_{U,142}$	$k_{F,DUB}$	$k_{U,DUB}$
125	3.1	25.8	614.2	43.8	11.0	524.0
147	5.2	23.0	301.3	38.8	12.0	1262.0
148	4.9	20.6	838.7	48.5	10.5	1018.7
149	4.4	20.8	175.9	21.2	11.3	928.4
151	3.6	15.8	279.5	30.6	8.4	392.1
152	5.4	17.1	275.1	43.1	9.7	94.9
153	5.1	14.2	280.6	40.6	9.8	377.0
154	3.2	11.4	142.7	142.7	39.5	498.4
156	4.4	26.0	449.3	44.3	13.0	809.0

Chapter 5 - Conclusion

With both structural and dynamic NMR experiments, this thesis has unequivocally verified that the LID domain of AK accesses a locally-unfolded intermediate conformation. By making entropy-enhancing, glycine mutations at surface-exposed residues this lab previously discovered this state in AK, however, it was not directly verified. The LU state is also functionally relevant as its population can modulate the observed binding affinity of the enzyme. Intermediate states in other proteins have been shown to be functionally important, however, most studies of intermediates have focused on structured states. The focus on structured intermediate states has been driven by reliance on crystallographic structures that can be visually inspected. These structures provide easily interpretable points of reference for the native and intermediate conformations and interpolations between structures are commonly used to describe how conformational changes occur. Structured intermediates are generally targeted by traditional mutations in the interior of the protein but interpretation of this mutational strategy is complicated since these mutations can also change the structures of the ground state and break interactions with other regions of the protein. Furthermore, structured states are not the only type of intermediates. The full conformational landscape and the contribution of intermediate states to function cannot be addressed without exploring additional kinds of intermediates.

The use of x-ray crystallography to identify intermediates necessarily restricts experimental focus to structured states and reliance on this technique can leave out other potentially important intermediates. Furthermore, mutation strategies that target unique structural intermediates leave other types of intermediates unexplored. Identifying and manipulating unfolded intermediates is more difficult because they lack a unique structure. Therefore, these states cannot be identified by crystallography and are generally unobservable by any technique given their low equilibrium populations. Traditional mutations that stabilize intermediate structures are not applicable in this case because there is no structure to stabilize. At best these mutations can be used to shift the equilibrium towards intermediate states by destabilizing the native state. However, an underappreciated feature of unfolded states is their degeneracy of conformations, i.e. many different conformations making up the unfolded state. This feature can be exploited by

increasing the number of unfolded conformations and shifting the equilibrium to favor unfolded states. The glycine mutation scheme used here targets the degeneracy of unfolded states by increasing the conformational entropy of the protein backbone. An added benefit of entropy-enhancing mutations is that the structure and stability of the native state remain largely unchanged, which simplifies interpretation of the results.

The available techniques to characterize unfolded states are limited due to the difficulties stated above. NMR is one of the few techniques available that can structurally characterize unfolding and provide information on a per-residue basis. Furthermore, NMR has the ability to directly interrogate conformation dynamics in proteins between different states. This is even possible for lowly-populated states, which is ideal for studying high-energy intermediates. Direct characterization of the LU intermediate in AK has become possible by combining NMR experiments with entropy-enhancing mutations that invert the equilibrium populations between the native and LU states. This has allowed chemical shift differences in HSQC spectra to be directly measured between the native and LU states in V142G and V135G/V142G, respectively. The chemical shift differences demonstrate a large structural rearrangement restricted to the LID domain with LID peaks collapsing from well-dispersed regions into the center of the amide proton spectrum. This result directly demonstrates a clear shift in population from a folded, native structure to a state with a locally-unfolded LID domain. Verified slow exchange pairs between the same two conformations in V135G provided further evidence that all variants access the same conformational transition.

Additionally, relaxation dispersion (RD) NMR experiments were used to investigate conformational dynamics between observed states and unobservable, lowly-populated states. These experiments also provide information about the populations, rates of exchange, and chemical shift differences between the exchanging states. However, a limitation of these experiments has been an inability to verify the structural differences between the observed and inferred states in all but a limited number of cases. The ability to invert the equilibrium from the folded, native state in V142G to the LU intermediate state in V135G/V142G has allowed investigation of the local unfolding conformational

transition using RD experiments from both sides of the equilibrium. The fact that the chemical shift differences from RD fits for both of these variants match one another provides strong evidence that both variant undergo the same conformational transition. More importantly, the agreement in chemical shift differences between RD fit parameters and directly measured quantities truly verifies these measurements. This has only been possible by rationally modulating the equilibrium populations with entropy-enhancing mutations.

The fact that the LU state is accessible in the native conformational landscape of AK and is functionally relevant requires a reassessment of the accepted mechanism for the function of this protein. The equilibrium populations between the open and closed states identified in crystal structures have been hypothesized by many as the main determinant of binding affinity in AK. As previous results in the lab have shown, this obviously not the case since the observed binding affinity can be modulated by the population of the LU state. Mutations have been made that modulate the equilibrium populations between the open and closed conformations of AK based on specific interactions observed in crystal structures. However, mutations that promote LID opening also break interactions with the AMPbd. Therefore, it is not clear if the LID domain is solely responsible for the observed behavior. Even more interesting is that single mutations in the LID domain that substantially increase the LU population yet have no effect on the catalytic rate of the enzyme. This result brings into question the hypothesis that opening of the LID domain determines the rate-limiting step of the enzymatic reaction. Experiments investigating local unfolding have not been able to measure opening and closing of the LID and AMPbd therefore this hypothesis cannot be directly tested. However, LID opening and local unfolding would have to be totally decoupled for this hypothesis to be true. Either unfolding has no effect on the opening rate of the LID or this hypothesis is false. Studies focusing on LID opening generally have not considered LID unfolding, however, a recent paper investigating urea induced enzymatic activation of AK required a three-state model including the LU intermediate to explain the data.[115]

The use of entropy-enhancing mutations that can rationally modulate the equilibrium populations of unfolded conformations opens the door to exploration of this understudied class of intermediate. As

stated above partially unfolded conformations are not likely to be identified using common techniques and traditional mutation strategies. Cases such as this study where unfolded intermediates have been identified have been through deliberate probing of the unfolded ensemble of conformations. The specific characteristics of unfolded intermediates and the dearth of studies specifically probing partially unfolded states leaves gaps in our understanding about the prevalence of local disorder and the role that disorder plays in function. This lab has pioneered a strategy to stabilize locally-unfolded intermediates while maintaining the original ground-state structure that allows exploration of a region of conformational space that is rarely investigated. Furthermore, the hypothesis that local unfolding contributes to conformational fluctuations between structured states suggests that local unfolding may be more widespread than previously thought. This hypothesis can be directly tested using entropy-enhancing mutations. Investigations of intermediate states have been confined to conformations that can be seen, resulting in a streetlight effect where only structural intermediates are explored. Analogously, we have developed a flashlight to look beyond the streetlight where hidden intermediate states that are functionally important lie.

References

1. Milo, R.; Phililips, R.; Orme, N., *Cell Biology by the numbers*. Garland Science: 2015.

2. Sumner, J. B., The Isolation and Crystallization of the Enzyme Urease. Preliminary paper. *Journal of Biological Chemistry* **1926,** *69* (2), 435-441.

3. Kendrew, J. C.; Bodo, G.; Dintzis, H. M.; Parrish, R. G.; Wyckoff, H.; Phillips, D. C., 3-Dimensional Model of the Myoglobin Molecule Obtained by X-Ray Analysis. *Nature* **1958,** *181* (4610), 662-666.

4. Perutz, M. F.; Rossmann, M. G.; Cullis, A. F.; Muirhead, H.; Will, G.; North, A. C. T., Structure of Haemoglobin - 3-Dimensional Fourier Synthesis at 5.5-a Resolution, Obtained by X-Ray Analysis. *Nature* **1960,** *185* (4711), 416-422.

5. Pauling, L.; Corey, R. B.; Branson, H. R., The Structure of Proteins - 2 Hydrogen-Bonded Helical Configurations of the Polypeptide Chain. *Proc. Natl. Acad. Sci. U. S. A.* **1951,** *37* (4), 205-211.

6. Anfinsen, C.; Haber, E., Studies on Reduction and Re-Formation of Protein Disulfide Bonds. *Journal of Biological Chemistry* **1961,** *236* (5), 1361-&.

7. Karplus, M.; McCammon, J. A., Molecular dynamics simulations of biomolecules. *Nat. Struct. Biol.* **2002,** *9* (9), 646-52.

8. Dill, K. A.; Ozkan, S. B.; Shell, M. S.; Weikl, T. R., The protein folding problem. *Annual Review of Biophysics* **2008,** *37*, 289-316.

9. Onuchic, J. N.; Wolynes, P. G., Theory of protein folding. *Curr Opin Struct Biol* **2004,** *14* (1), 70-5.

10. Hartl, F. U.; Hayer-Hartl, M., Converging concepts of protein folding in vitro and in vivo. *Nat. Struct. Biol.* **2009,** *16* (6), 574-81.

11. Schellman, J. A., Destabilization and stabilization of proteins. *Q Rev Biophys* **2005,** *38* (4), 351-61.

12. Whitford, P. C.; Sanbonmatsu, K. Y.; Onuchic, J. N., Biomolecular dynamics: order-disorder transitions and energy landscapes. *Rep. Prog. Phys.* **2012,** *75* (7), 076601.

13. Bryngelson, J. D.; Wolynes, P. G., Intermediates and Barrier Crossing in a Random Energy-Model (with Applications to Protein Folding). *J. Phys. Chem.* **1989,** *93* (19), 6902-6915.

14. Leopold, P. E.; Montal, M.; Onuchic, J. N., Protein Folding Funnels - a Kinetic Approach to the Sequence Structure Relationship. *Proc. Natl. Acad. Sci. U. S. A.* **1992,** *89* (18), 8721-8725.

15. Englander, S. W.; Sosnick, T. R.; Mayne, L.; Bai, Y., Hydrogen exchange: the modern legacy of Linderstrom-Lang. **1997**.

16. Hu, W.; Kan, Z. Y.; Mayne, L.; Englander, S. W., Cytochrome c folds through foldon-dependent native-like intermediates in an ordered pathway. *Proc Natl Acad Sci U S A* **2016,** *113* (14), 3809-14.

17. Koshland, D. E., Enzyme Flexibility and Enzyme Action. *J. Cell. Physiol.* **1959,** *54* (S1), 13.

18. Perutz, M. F., Haem-Haem Interaction and the Problem of Allostery. *Nature* **1970,** *228*.

19. Motlagh, H. N.; Wrabl, J. O.; Li, J.; Hilser, V. J., The ensemble nature of allostery. *Nature* **2014,** *508* (7496), 331-9.

20. Tzeng, S. R.; Kalodimos, C. G., Allosteric inhibition through suppression of transient conformational states. *Nat Chem Biol* **2013,** *9* (7), 462-5.

21. Beece, D.; Eisenstein, L.; Frauenfelder, H.; Good, D.; Marden, M. C.; Reinisch, L.; Reynolds, A. H.; Sorensen, L. B.; Yue, K. T., Solvent viscosity and protein dynamics. *Biochemistry* **1980,** *19* (23), 5147-57.

22. Perutz, M. F.; Wilkinson, A. J.; Paoli, M.; Dodson, G. G., The stereochemical mechanism of the cooperative effects in hemoglobin revisited. *Annu. Rev. Biophys. Biomol. Struct.* **1998,** *27*, 1-34.

23. Hilser, V. J.; Anderson, J. A.; Motlagh, H. N., Allostery vs. "allokairy". *Proc Natl Acad Sci U S A* **2015,** *112* (37), 11430-1.

24. Motlagh, H. N.; Hilser, V. J., Agonism/antagonism switching in allosteric ensembles. *Proc Natl Acad Sci U S A* **2012,** *109* (11), 4134-9.

25. Cooper, A.; Dryden, D. T. F., Allostery without Conformational Change - a Plausible Model. *Eur Biophys J Biophy* **1984,** *11* (2), 103-109.

26. Wright, P. E.; Dyson, H. J., Intrinsically Unstructured Proteins: Re-assessing the Protein Structure-Function Paradigm. *J Mol Biol* **1999,** (293), 293-321.

27. Olsson, M. H. M.; Parson, W. W.; Warshel, A., Dynamical contributions to enzyme catalysis: Critical tests of a popular hypothesis. *Chem. Rev.* **2006,** *106* (5), 1737-1756.

28. Allemann, R. K., Unraveling the role of protein dynamics in dihydrofolate reductase catalysis. *Proc Natl Acad Sci U S A* **2013,** *110* (41).

29. Hammes-Schiffer, S.; Benkovic, S. J., Relating Protein Motion to Catalysis. *Annu Rev Biochem* **2006,** *75*.

30. Schulenburg, C.; Hilvert, D., Protein conformational disorder and enzyme catalysis. *Top Curr Chem* **2013,** *337*, 41-67.

31. Klinman, J. P., Dynamically achieved active site precision in enzyme catalysis. *Acc Chem Res* **2015,** *48* (2), 449-56.

32. Frushicheva, M. P.; Mills, M. J.; Schopf, P.; Singh, M. K.; Prasad, R. B.; Warshel, A., Computer aided enzyme design and catalytic concepts. *Curr. Opin. Chem. Biol.* **2014,** *21*, 56-62.

33. Whittington, A. C.; Larion, M.; Bowler, J. M.; Ramsey, K. M.; Bruschweiler, R.; Miller, B. G., Dual allosteric activation mechanisms in monomeric human glucokinase. *Proc Natl Acad Sci U S A* **2015,** *112* (37), 11553-8.

34. Buller, A. R.; Brinkmann-Chen, S.; Romney, D. K.; Herger, M.; Murciano-Calles, J.; Arnold, F. H., Directed evolution of the tryptophan synthase beta-subunit for stand-alone function recapitulates allosteric activation. *Proc Natl Acad Sci U S A* **2015,** *112* (47), 14599-604.

35. Campbell, E.; Kaltenbach, M.; Correy, G. J.; Carr, P. D.; Porebski, B. T.; Livingstone, E. K.; Afriat-Jurnou, L.; Buckle, A. M.; Weik, M.; Hollfelder, F.; Tokuriki, N.; Jackson, C. J., The role of protein dynamics in the evolution of new enzyme function. *Nat Chem Biol* **2016**.

36. Boehr, D. D.; McElheny, D.; Dyson, H. J.; Wright, P. E., The dynamic energy landscape of dihydrofolate reductase catalysis. *Science* **2006,** *313* (5793), 1638-42.

37. Somero, G. N.; Fields, P. A., Hot spots in cold adaptations: Localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc Natl Acad Sci U S A* **1998,** *95*.

38. Feller, G., Psychrophilic enzymes: from folding to function and biotechnology. *Scientifica (Cairo)* **2013,** *2013*, 512840.

39. D'Amico, S.; Claverie, P.; Collins, T.; Georlette, D.; Gratia, E.; Hoyoux, A.; Meuwis, M. A.; Feller, G.; Gerday, C., Molecular basis of cold adaptation. *Philos Trans R Soc Lond B Biol Sci* **2002,** *357* (1423), 917-25.

40. Feller, G.; Gerday, C., Psychrophilic enzymes: hot topics in cold adaptation. *Nat Rev Microbiol* **2003,** *1* (3), 200-8.

41. Fields, P. A.; Dong, Y.; Meng, X.; Somero, G. N., Adaptations of protein structure and function to temperature: there is more than one way to 'skin a cat'. *J Exp Biol* **2015,** *218* (Pt 12), 1801-11.

42. Somero, G. N., Temperature Adaptation of Enzymes: Biological Optimization Through Structure-Function Compromises. *Ann. Rev. Ecol. Syst.* **1978,** *9*.

43. Elias, M.; Wieczorek, G.; Rosenne, S.; Tawfik, D. S., The universality of enzymatic rate-temperature dependency. *Trends Biochem. Sci.* **2014,** *39* (1), 1-7.

44. Zavodszky, P.; Kardos, J.; Svingor, A.; Petsko, G. A., Adjustment of conformational flexibility is a key event in the thermal adaptation of proteins. *Proc. Natl. Acad. Sci. U. S. A.* **1998,** *95* (13), 7406-7411.

45. Fields, P. A.; Somero, G. N., Hot spots in cold adaptation: Localized increases in conformational flexibility in lactate dehydrogenase A4 orthologs of Antarctic notothenioid fishes. *Proc Natl Acad Sci U S A* **1998,** *95*.

46. Schulz, G. E.; Muller, C. W., Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure* **1996,** *4* (2).

47. Muller, C. W.; Schlauderer, G. J.; Reinstein, J.; Schulz, G. E., Adenylate kinase motions during catalysis: an energetic counterweight balancing substrate binding. *Structure* **1996,** *4* (2), 147-56.

48. Wolf-Watz, M.; Thai, V.; Henzler-Wildman, K.; Hadjipavlou, G.; Eisenmesser, E. Z.; Kern, D., Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair. *Nat Struct Mol Biol* **2004,** *11* (10), 945-9.

49. Bae, E.; Phillips, G. N., Jr., Roles of static and dynamic domains in stability and catalysis of adenylate kinase. *Proc Natl Acad Sci U S A* **2006,** *103* (7), 2132-7.

50. Henzler-Wildman, K. A.; Lei, M.; Thai, V.; Kerns, S. J.; Karplus, M.; Kern, D., A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **2007,** *450* (7171), 913-6.

51. Henzler-Wildman, K. A.; Thai, V.; Lei, M.; Ott, M.; Wolf-Watz, M.; Fenn, T.; Pozharski, E.; Wilson, M. A.; Petsko, G. A.; Karplus, M.; Hubner, C. G.; Kern, D., Intrinsic motions along an enzymatic reaction trajectory. *Nature* **2007,** *450* (7171), 838-44.

52. Pelz, B.; Zoldak, G.; Zeller, F.; Zacharias, M.; Rief, M., Subnanometre enzyme mechanics probed by single-molecule force spectroscopy. *Nat Commun* **2016,** *7*, 10848.

53. Kovermann, M.; Aden, J.; Grundstrom, C.; Elisabeth Sauer-Eriksson, A.; Sauer, U. H.; Wolf-Watz, M., Structural basis for catalytically restrictive dynamics of a high-energy enzyme state. *Nat Commun* **2015,** *6*, 7644.

54. Kerns, S. J.; Agafonov, R. V.; Cho, Y. J.; Pontiggia, F.; Otten, R.; Pachov, D. V.; Kutter, S.; Phung, L. A.; Murphy, P. N.; Thai, V.; Alber, T.; Hagan, M. F.; Kern, D., The energy landscape of adenylate kinase during catalysis. *Nat Struct Mol Biol* **2015,** *22* (2), 124-31.

55. Beckstein, O.; Denning, E. J.; Perilla, J. R.; Woolf, T. B., Zipping and unzipping of adenylate kinase: atomistic insights into the ensemble of open<-->closed transitions. *J Mol Biol* **2009,** *394* (1), 160-76.

56. Daily, M. D.; Makowski, L.; Phillips, G. N., Jr.; Cui, Q., Large-scale motions in the adenylate kinase solution ensemble: coarse-grained simulations and comparison with solution X-ray scattering. *Chem. Phys.* **2012,** *396*, 84-91.

57. Brokaw, J. B.; Chu, J. W., On the roles of substrate binding and hinge unfolding in conformational changes of adenylate kinase. *Biophys J* **2010,** *99* (10), 3420-9.

58. Hanson, J. A.; Duderstadt, K.; Watkins, L. P.; Bhattacharyya, S.; Brokaw, J.; Chu, J. W.; Yang, H., Illuminating the mechanistic roles of enzyme conformational dynamics. *Proc Natl Acad Sci U S A* **2007,** *104* (46), 18055-60.

59. Song, H. D.; Zhu, F., Conformational dynamics of a ligand-free adenylate kinase. *PLoS One* **2013,** *8* (7), e68023.

60. Matsunaga, Y.; Fujisaki, H.; Terada, T.; Furuta, T.; Moritsugu, K.; Kidera, A., Minimum free energy path of ligand-induced transition in adenylate kinase. *PLoS Comput Biol* **2012,** *8* (6), e1002555.

61. Zeller, F.; Zacharias, M., Substrate Binding Specifically Modulates Domain Arrangements in Adenylate Kinase. *Biophys J* **2015,** *109* (9), 1978-85.

62. Jana, B.; Adkar, B. V.; Biswas, R.; Bagchi, B., Dynamic coupling between the LID and NMP domain motions in the catalytic conversion of ATP and AMP to ADP by adenylate kinase. *J Chem Phys* **2011,** *134* (3), 035101.

63. Esteban-Martin, S.; Fenwick, R. B.; Aden, J.; Cossins, B.; Bertoncini, C. W.; Guallar, V.; Wolf-Watz, M.; Salvatella, X., Correlated inter-domain motions in adenylate kinase. *PLoS Comput Biol* **2014,** *10* (7), e1003721.

64. Unan, H.; Yildirim, A.; Tekpinar, M., Opening mechanism of adenylate kinase can vary according to selected molecular dynamics force field. *J Comput Aided Mol Des* **2015,** *29* (7), 655-65.

65. Fujii, A.; Sekiguchi, Y.; Matsumura, H.; Inoue, T.; Chung, W. S.; Hirota, S.; Matsuo, T., Excimer emission properties on pyrene-labeled protein surface: correlation between emission spectra, ring stacking modes, and flexibilities of pyrene probes. *Bioconjug. Chem.* **2015,** *26* (3), 537-48.

66. Nagarajan, S.; Amir, D.; Grupi, A.; Goldenberg, D. P.; Minton, A. P.; Haas, E., Modulation of functionally significant conformational equilibria in adenylate kinase by high concentrations of trimethylamine oxide attributed to volume exclusion. *Biophys J* **2011,** *100* (12), 2991-9.

67. Aden, J.; Verma, A.; Schug, A.; Wolf-Watz, M., Modulation of a pre-existing conformational equilibrium tunes adenylate kinase activity. *J Am Chem Soc* **2012,** *134* (40), 16562-70.

68. Perkins, S. J.; Wüthrich, K., Ring Current Effects in the Conformation Dependent NMR Chemical-Shifts of Aliphatic Protons in the Basic Pancreatic Trypsin-Inhibitor. *Biochim. Biophys. Acta* **1979,** *576* (2), 409-423.

69. Whitford, P. C.; Miyashita, O.; Levy, Y.; Onuchic, J. N., Conformational transitions of adenylate kinase: switching by cracking. *J Mol Biol* **2007,** *366* (5), 1661-71.

70. Berry, M. B.; Phillips, G. N., Jr., Crystal structures of Bacillus stearothermophilus adenylate kinase with bound Ap5A, Mg^{2+} Ap5A, and Mn^{2+} Ap5A reveal an intermediate lid position and six coordinate octahedral geometry for bound Mg^{2+} and Mn^{2+}. *Proteins* **1998,** *32* (3), 276-88.

71. Hilser, V. J.; Freire, E., Structure-based Calculation of the Equilibruim Folding Pathway of Proteins. Correlation with Hydrogen Exchange Protection Factors. *J Mol Biol* **1996,** *262*.

72. Gu, J.; Hilser, V. J., Predicting the energetics of conformational fluctuations in proteins from sequence: a strategy for profiling the proteome. *Structure* **2008,** *16* (11), 1627-37.

73. Englander, S. W.; Mayne, L.; Krishna, M. M., Protein folding and misfolding: mechanism and principles. *Q Rev Biophys* **2007,** *40* (4), 287-326.

74. Skinner, J. J.; Lim, W. K.; Bedard, S.; Black, B. E.; Englander, S. W., Protein dynamics viewed by hydrogen exchange. *Protein Sci* **2012,** *21* (7), 996-1005.

75. Lee, K. H.; Xie, D.; Freire, E.; Amzel, L. M., Estimation of changes in side chain configurational entropy in binding and folding: general methods and application to helix formation. *Proteins* **1994,** *20* (1), 68-84.

76. D'Aquino, J. A.; Gomez, J.; Hilser, V. J.; Lee, K. H.; Amzel, L. M.; Freire, E., The magnitude of the backbone conformational entropy change in protein folding. *Proteins* **1996,** *25* (2), 143-56.

77. Scott, K. A.; Alonso, D. O. V.; Sato, S.; Fersht, A. R.; Daggett, V., Conformational entropy of alanine versus glycine in protein denatured states. *Proc. Natl. Acad. Sci. U. S. A.* **2007,** *104* (8), 2661-2666.

78. Ferreon, J. C.; Hilser, V. J., The effect of the polyproline II (PPII) conformation on the denatured state entropy. *Protein Science* **2003,** *12* (3), 447-457.

79. Ferreon, J. C.; Hilser, V. J., Ligand-induced changes in dynamics in the RT loop of the C-terminal SH3 domain of Sem-5 indicate cooperative conformational coupling. *Protein Sci* **2003,** *12* (5), 982-96.

80. Parra, R. G.; Schafer, N. P.; Radusky, L. G.; Tsai, M. Y.; Guzovsky, A. B.; Wolynes, P. G.; Ferreiro, D. U., Protein Frustratometer 2: a tool to localize energetic frustration in protein molecules, now with electrostatics. *Nucleic Acids Res* **2016,** *44* (W1), W356-60.

81. Schrank, T. P.; Bolen, D. W.; Hilser, V. J., Rational modulation of conformational fluctuations in adenylate kinase reveals a local unfolding mechanism for allostery and functional adaptation in proteins. *Proc Natl Acad Sci U S A* **2009,** *106* (40), 16984-9.

82. Schrank, T. P. Investigations of Local Unfolding Like Conformational Excursions in the Native State of Adenylate Kinase. University of Texas Graduate School of Biomedical Sciences at Galveston, 2010.

83. Schrank, T. P.; Elam, W. A.; Li, J.; Hilser, V. J., Strategies for the thermodynamic characterization of linked binding/local folding reactions within the native state application to the LID domain of adenylate kinase from Escherichia coli. *Methods Enzymol* **2011,** *492*, 253-82.

84. Hansen, D. F.; Vallurupalli, P.; Kay, L. E., Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J Biomol NMR* **2008,** *41* (3), 113-20.

85. Sekhar, A.; Rumfeldt, J. A. O.; Broom, H. R.; Doyle, C. M.; Sobering, R. E.; Meiering, E. M.; Kay, L. E., Probing the free energy landscapes of ALS disease mutants of SOD1 by NMR spectroscopy. *Proc. Natl. Acad. Sci. U. S. A.* **2016**, 201611418.

86. Kay, L. E.; Hansen, D. F., Selective Characterization of Microsecond Motions in Proteins by NMR Relaxation. *J. Am. Chem. Soc.* **2009,** *2009* (131).

87. O'Connell, N. E.; Grey, M. J.; Tang, Y.; Kosuri, P.; Miloushev, V. Z.; Raleigh, D. P.; Palmer, A. G., 3rd, Partially folded equilibrium intermediate of the villin headpiece HP67 defined by ^{13}C relaxation dispersion. *J Biomol NMR* **2009,** *45* (1-2), 85-98.

88. Grey, M. J.; Wang, C. Y.; Palmer, A. G., Disulfide bond isomerization in basic pancreatic trypsin inhibitor: Multisite chemical exchange quantified by CPMG relaxation dispersion and chemical shift modeling. *J. Am. Chem. Soc.* **2003,** *125* (47), 14324-14335.

89. Zhang, Y.; Kitazawa, S.; Peran, I.; Stenzoski, N.; McCallum, S. A.; Raleigh, D. P.; Royer, C. A., High Pressure ZZ-Exchange NMR Reveals Key Features of Protein Folding Transition States. *J Am Chem Soc* **2016**.

90. Schneider, R.; Maurin, D.; Communie, G.; Kragelj, J.; Hansen, D. F.; Ruigrok, R. W.; Jensen, M. R.; Blackledge, M., Visualizing the molecular recognition trajectory of an intrinsically disordered protein using multinuclear relaxation dispersion NMR. *J Am Chem Soc* **2015,** *137* (3), 1220-9.

91. Anthis, N. J.; Clore, G. M., Visualizing transient dark states by NMR spectroscopy. *Q Rev Biophys* **2015,** *48* (1), 35-116.

92. Schulz, G. E.; Muller, C. W., Structure of the Complex Between Adenylate Kinase form Escherichia coli and the Inhibitor Ap5A Refined at 1.9 A Resolution; A Model for a Catalytic Transition State. *J. Mol. Biol.* **1992,** *224*.

93. Rule, G. S.; Hitchens, T. K., *Fundamentals of protein NMR spectroscopy*. Springer: Dordrecht, 2006; 530 p.

94. Williamson, M. P., Using chemical shift perturbation to characterise ligand binding. *Prog Nucl Magn Reson Spectrosc* **2013,** *73*, 1-16.

95. Murray, V.; Huang, Y.; Chen, J.; Wang, J.; Li, Q., A novel bacterial expression method with optimized parameters for very high yield production of triple-labeled proteins. *Methods Mol Biol* **2012,** *831*, 1-18.

96. Pervushin, K.; Riek, R.; Wider, G.; Wuthrich, K., Attenuated T_2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc. Natl. Acad. Sci. U. S. A.* **1997,** *94* (23), 12366-12371.

97. Skinner, S. P.; Fogh, R. H.; Boucher, W.; Ragan, T. J.; Mureddu, L. G.; Vuister, G. W., CCPNMR AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR* **2016**.

98. Cavanagh, J., *Protein NMR spectroscopy : principles and practice*. 2nd ed.; Academic Press: Amsterdam ; Boston, 2007; p xxv, 885 p.

99. Wishart, D. S.; Bigam, C. G.; Holm, A.; Hodges, R. S.; Sykes, B. D., H-1, C-13 and N-15 Random Coil Nmr Chemical-Shifts of the Common Amino-Acids .1. Investigations of Nearest-Neighbor Effects (Vol 5, Pg 67, 1995). *J. Biomol. NMR* **1995,** *5* (3), 332-332.

100. Li, W.; Wolynes, P. G.; Takada, S., Frustration, specific sequence dependence, and nonlinearity in large-amplitude fluctuations of allosteric proteins. *Proc Natl Acad Sci U S A* **2011,** *108* (9), 3504-9.

101. Cho, S. S.; Levy, Y.; Onuchic, J. N.; Wolynes, P. G., Overcoming residual frustration in domain-swapping: the roles of disulfide bonds in dimerization and aggregation. *Phys. Biol.* **2005,** *2* (2), S44-55.

102. Baldwin, A. J.; Kay, L. E., NMR spectroscopy brings invisible protein states into focus. *Nat Chem Biol* **2009,** *5* (11), 808-14.

103. Richards, R. E.; Carver, J. P., A General Two-Site Solution for the Chemical Exchange Produced Dependence of T_2 Upon the Carr-Purcel Pulse Separation. *J. Magn. Reson.* **1972,** *6*.

104. Hansen, D. F.; Vallurupalli, P.; Kay, L. E., An improved 15N relaxation dispersion experiment for the measurement of millisecond time-scale dynamics in proteins. *J Phys Chem B* **2008,** *112* (19), 5898-904.

105. d'Auvergne, E. J.; Gooley, P. R., Optimisation of NMR dynamic models I. Minimisation algorithms and their performance within the model-free and Brownian rotational diffusion spaces. *J Biomol NMR* **2008,** *40* (2), 107-19.

106. Kay, L. E.; Forman-Kay, J. D.; Farrow, N. A.; Zhang, O., A heteronuclear correlation experiment for simultaneous determination of ^{15}N longitudenal decay and chemical exchange rates of systems in slow equilibrium. *J. Biomol. NMR* **1994,** *4*.

107. Kloiber, K.; Spitzer, R.; Grutsch, S.; Kreutz, C.; Tollinger, M., Longitudinal exchange: an alternative strategy towards quantification of dynamics parameters in ZZ exchange spectroscopy. *J Biomol NMR* **2011,** *51* (1-2), 123-9.

108. Tollinger, M.; Skrynnikov, N. R.; Mulder, F. A.; Forman-Kay, J. D.; Kay, L. E., Slow dynamics in folded and unfolded states of an SH3 domain. *J Am Chem Soc* **2001,** *123* (46), 11341-52.

109. Ishima, R.; Torchia, D. A., Estimating the time scale of chemical exchange of proteins from measurements of transverse relaxation rates in solution. *J. Biomol. NMR* **1999,** *14* (4), 369-372.

110. Ding, K.; Louis, J. M.; Gronenborn, A. M., Insights into Conformation and Dynamics of Protein GB1 During Folding and Unfolding by NMR. *Journal of Molecular Biology* **2004,** *335* (5), 1299-1307.

111. Neudecker, P.; Zarrine-Afsar, A.; Davidson, A. R.; Kay, L. E., Phi-value analysis of a three-state protein folding pathway by NMR relaxation dispersion spectroscopy. *Proc Natl Acad Sci U S A* **2007,** *104* (40), 15717-22.

112. Okuda, M.; Higo, J.; Komatsu, T.; Konuma, T.; Sugase, K.; Nishimura, Y., Dynamics of the Extended String-Like Interaction of TFIIE with the p62 Subunit of TFIIH. *Biophys J* **2016,** *111* (5), 950-62.

113. Korzhnev, D. M.; Neudecker, P.; Mittermaier, A.; Orekhov, V. Y.; Kay, L. E., Multiple-site exchange in proteins studied with a suite of six NMR relaxation dispersion experiments: An application to the folding of a Fyn SH3 domain mutant. *J. Am. Chem. Soc.* **2005,** *127* (44), 15602-15611.

114. Ivankov, D. N.; Garbuzynskiy, S. O.; Alm, E.; Plaxco, K. W.; Baker, D.; Finkelstein, A. V., Contact order revisited: influence of protein size on the folding rate. *Protein Sci* **2003,** *12* (9), 2057-62.

115. Rogne, P.; Wolf-Watz, M., Urea-Dependent Adenylate Kinase Activation following Redistribution of Structural States. *Biophys J* **2016,** *111* (7), 1385-1395.

Vita

Jeremy A. Anderson was born in Marshfield, Wisconsin on November 23rd, 1987 and grew up in the small North Central Wisconsin community of Owen-Withee. He attended the University of Minnesota, Duluth and conducted endocrinology research under the guidance of Dr. Grant W. Anderson (no relation). Jeremy graduated with a Bachelor of Science degree in Biochemistry and Molecular Biology, a Bachelor of Arts degree in Chemistry, and a minor in Physics. Upon graduating Jeremy rode his bicycle solo 1000 miles from Marshfield, WI to Bowling Green, KY to visit distant relatives in La Vergne, TN and hitched a ride back to Wisconsin on the tour bus of country music band Sawyer Brown. The following fall Jeremy began his graduate work in the Program in Molecular Biophysics at The Johns Hopkins University in Baltimore, Maryland. Jeremy sought to combine his interests in biology and biochemistry with the quantitative nature of physics and did so under the mentorship of Vincent. J. Hilser. Upon completion of his thesis Jeremy will pursue a career in academia beginning with a postdoctoral research position in the laboratory of Michael J. Harms at the University of Oregon in Eugene.